

| | | | |
|---|--|--|-----------------------------------|
| DR | HSP: | P02699; | 1F88. |
| DR | InterPro: | IPR000276; | GPCR_Rhodpsn. |
| DR | Pfam: | PF00001; | 7tm_1; 1. |
| DR | PRINTS: | PR00237; | GPCRRHODOPSN. |
| DR | PROSITE: | PS00237; | G_PROTEIN_RECEP_F1_1; UNKNOWN_1. |
| DR | PROSITE: | PS00262; | G_PROTEIN_RECEP_F1_2; 1. |
| KW | Receptor. | | |
| SQ | SEQUENCE | 447 AA; | 48418 MW; E3DAAE5EE1F0FB99 CRC64; |
| | | | |
| Query Match 9.9%; Score 217; DB 4; Length 447; | | | |
| Best Local Similarity 22.5%; Pred.No. 6.6e-11; | | | |
| Matches 89; Conservative 65; Mismatches 130; Indels 112; Gaps | | | |
| | | | |
| QY | 36 | YPLSFQVSTGFLMLIEVLGLGSNLTVLVLCMKGNLINSVSNITMNLHLVDIVICVC | 95 |
| | | : : : : : : : : : : : : : : : : | |
| DB | 61 | YAVLFMSVGGNNLIIVLGL-----SRRLRTVTNAFLLSLAVIDLLAVAC | 107 |
| | | | |
| QY | 96 | IPIITVILLLESNTALICCPEHCVSF-----ASVSTAINFAITLDKYDISVKP--A | 148 |
| | | : : : : : : : : : : : : : : : : | |
| DB | 108 | MPTETLPNLMGTGTFEFTVIC-----KAVSYLMGVSVSVST-LSLVAIALERYSAICRPLQA | 162 |
| | | | |
| QY | 149 | NRIILTWGRAVMIAMISWIFFSFSLPIFLEVNFSS-----LQSGTNWKNTLLCVSTNE | 202 |
| | | : : : : : : : : : : : : : : : | |
| DB | 163 | RVMQTRSHAARVIAVWLUSGL-LMPVPVVYTVVPGVRVLCQCHRWPS----- | 211 |
| | | | |
| QY | 203 | YYTELGMYYHLVQLQIFIFFFTVVMIIYTKILOALNIRIGTRF----- | 246 |
| | | : : : : : : : : : : : : : : : | |
| DB | 212 | --ARVQTWSVLLLLLLLFFIPGVMAVAGLLISREL--YLGLRFDGSDSDSQSRVRNQ | 267 |
| | | | |
| QY | 247 | -----STGQKKK-----ARKKTISLTTOHEATDMQSQSGGRNVVFG | 283 |
| | | : : : : : : : : : : : : : : | |
| DB | 268 | GLPCGAHQNGRCRPETGAVGKDSDGCYVOLPRSPRALEITA---LTAPGPGSGSR---- | 319 |
| | | | |
| QY | 284 | VRTSVSVIITALRAVHRERERRQRKRFRMSLLIISIFLLCWTTPISVLNTTILCLGPS- | 342 |
| | | : : : : : : : : : : : : : : : | |
| DB | 320 | -PTQAKLL-----AKRVVMULLVTIVFLFCVLPVYSANTWRADFPGA | 363 |
| | | | |
| QY | 343 | -DLLVLRLLCF-LVMAYGTTIFPHLYAFTRFKOK 376 | |
| | | : : : : : : : : : : : : : : | |
| DB | 364 | HRALSGAPTISFIHLISYASACVNPLYCEPMHRRFRQ 399 | |
| | | | |
| RESULT 14 | | | |
| Q98843 | ID | Q98843 | PRELIMINARY; PRT; 458 AA. |
| AC | Q98843 | | |
| DT | 01-FEB-1997 | (TREMBLrel. 02, Created) | |
| DT | 01-FEB-1997 | (TREMBLrel. 02, Last sequence update) | |
| DT | 01-MAR-2002 | (TREMBLrel. 20, Last annotation update) | |
| DE | Dopamine D1B receptor. | | |
| DN | D1B. | | |
| OS | Anguilla anguilla (European freshwater eel). | | |
| OC | Eukaryota; Metazoa; Chordata; Teleostei; Vertebrata; Euteleostomi; | | |
| OC | Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae; | | |
| OC | Anguilla. | | |
| OX | NCBI_TaxID=7936; | | |
| FP | [1] | | |
| RN | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=97160583; PubMed=9006917; | | |
| RA | Cardinaud B., Sugamori K.S., Coudouet S., Vincent J.D., Niznik H.B., | | |
| RA | Vernier P.; | | |
| RT | "Early emergence of three dopamine D1 receptor subtypes in | | |
| RT | vertebrates. Molecular phylogenetic, pharmacological, and functional | | |
| RT | criteria defining D1A, D1B, and D1C receptors in European eel Anguilla | | |
| RT | anguilla."; | | |
| RL | J. Biol. Chem. 272:2778-2787(1997). | | |
| CC | -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY). | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | |
| ENBL | U62920; AAC50070.1; - | | |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. | | |
| DR | PRINTS: PR00237; GPCRHHODOPSN. | | |
| DR | PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1. | | |

| | | |
|---|--|---------------------------------|
| RL | Cytogenet. Cell Genet. | 65:184-185(1994). |
| [3] | | |
| RN | SEQUENCE FROM N.A. | |
| RP | TISSUE=LUNG; | |
| RC | MEDLINE=95151633; PubMed=7848914; | |
| RX | Ito M., Iwata N., Taniguchi T., Murayama T., Chihara K., Matsui T.; | |
| RA | "Functional characterization of two cholecystokinin-B/gastrin receptor | |
| RT | isoforms: a preferential splice donor site in the human receptor | |
| RF | gene."; | |
| RL | Cell Growth Differ. | 5:1127-1135(1994). |
| RT | | |
| RN | SEQUENCE OF 16-26 FROM N.A. | |
| RP | MEDLINE=95194412; PubMed=7887934; | |
| RX | Miyake A.; | |
| RA | "A truncated isoform of human CCK-B/gastrin receptor generated by | |
| RT | alternative usage of a novel exon."; | |
| EL | Biochem. Biophys. Res. Commun. | 208:230-237(1995). |
| DR | EMBL; D21219; BAA04759.1; -; | |
| DR | EMBL; S76072; AAB33740.1; -; | |
| DR | HSPB; P02699; IF88. | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | |
| DR | Pfam; PF00001; 7tm_1; 1. | |
| DR | PRINTS; PR00237; GPCRRHODPSN. | |
| DR | PROSITE; PS00237; G-PROTEIN_RECEPT_F1_1; UNKNOWN_1. | |
| DR | PROSITE; PSS0262; G-PROTEIN_RECEP_F1_2; 1. | |
| FT | NON_TER | 396 |
| SQ | SEQUENCE | 396 AA; A25DD5597CA903C0 CRC64; |
| | | |
| Query Match | 9.9%; Score 218; DB 4; Length 396; | |
| Best Local Similarity | 22.8%; Pred. No. 4.Be-11; | |
| Matches | 88; Conservative 70; Mismatches 136; Indels 92; Gaps 15; | |
| | | |
| QY | 36 YPLSFQVSLGFLMLEIVLGLSGNLTVLVLCMKSNLSNVSIITMNLHLVDLIICVGC 95 | |
| DB | 10 YAVIFLMSVGGNMLIIYVLGL-----SRRLRVTNAFLSLAVSDLLLAAC 56 | |
| | | |
| QY | 96 IPLRVTIILLLESNTALICCFHEACYSF-----ASYSTAINVFATLDRIYDISVKP--A 148 | |
| DB | 57 MPFTLLPNLMGTFFIGTVIC---KAVSYLMGVSVSVYST-LSLVAIALERYSAICRPQA 111 | |
| QY | 149 NRILTMGRVAMLMISWIFSFLIEFIEVNFFS-----LQSGNTWENKTLLCVSTNE 202 | |
| DB | 112 RVWQTRSHAARVIATWLLLSGL-LMVPPYPVVTVQPVPGRVLOCVHRWS----- 160 | |
| QY | 203 YYTELGMVYHLLVOIPFFEFFVVMYLITYTKILOALNRIGTTF-----STGOKKKARKK 257 | |
| DB | 161 --ARVRTQSWVLLLLLLLFFFPVGVMAVAYGLISREL--YLGRLFDGSDSDSQSRVRNQG 216 | |
| QY | 258 KTISLTQH-----EATDMSSQSGGRNVVF-----GVRTSVSVIIA 293 | |
| DB | 217 GLPGAVHQNGRCRPTGAVGSDGCVQLPRSPALELTALTAPGPGSGRPTQAKLLA 276 | |
| QY | 294 LRRVAKHRRERERQKRFRNSLLIIITFLCLWTPISVNLTTILCGPS--DLLVLKRLC 351 | |
| DB | 277 -----KKRVRMLLVIVVLFCLWLPVYSANTWRAFDFGGAHRSALSGAPIS 322 | |
| QY | 352 FL-VMAVGYTTIFFPLLVAFTROKFQK 376 | |
| DB | 323 FIIHLSVASACVNPVLCFMHRRFRQ 348 | |
| | | |
| RESULT 12 | | |
| Q9N296 | PRELIMINARY; | PRT; 422 AA. |
| ID Q9N296 | | |
| AC Q9N296; | | |
| DT 01-OCT-2000 (TrEMBLrel. 15, Created) | | |
| DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update) | | |
| DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update) | | |
| DE Serotonin receptor 1A. | | |
| GN HRI1A. | | |
| OS Pongo pygmaeus (Orangutan). | | |
| OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; | | |
| OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pongo. | | |

```
RN SEQUENCE FROM N.A.
RX MEDLINE=98068842; PubMed=9407007;
RA Lundell I., Berglund M.M., Starback P., Salaneck E., Gehlert D.R.,
RA Larhammar D.;
RT "Cloning and characterization of a novel neuropeptide Y receptor
RT subtype in the zebrafish.";
RL DNA Cell Biol. 16:1357-1363(1997).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=98096393; PubMed=9434780;
RA Ringvall M., Berglund M.M., Larhammar D.;
RT "Multiplicity of neuropeptide Y receptors: cloning of a third distinct
RT subtype in the zebrafish.";
RL Biochem. Biophys. Res. Commun. 241:749-755(1997).
DR EMBL; AF037401; AAC41277.1; -.
DR ZFIN; ZDB-GENE-990415-175; nbyryc.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 373 AA; 41673 MW; 217DA7F51A940CFD CRC64;

Query Match 10.1%; Score 222; DB 13; Length 373;
Best Local Similarity 21.8%; Pred. No. 2e-11;
Matches 90; Conservative 75; Mismatches 123; Indels 124; Gaps 19;

Qy 14 ESNITVRDDIDD-----INTNMYQPLSYPLSFQVSLTGFLMLE-----IVLGLGSNLTV 62
Db 1 ESNITVRDDIDD-----INTNMYQPLSYPLSFQVSLTGFLMLE-----IVLGLGSNLTV 62
Qy 63 LVLYCMKSNLINSVNIITMNLHVLDVLCVGCIPLTIVILL-----SLESNTALIC 115
Db 1 LVLYCMKSNLINSVNIITMNLHVLDVLCVGCIPLTIVILL-----SLESNTALIC 115
Qy 51 LVFIISQKEMNTVNIILANLSCSDILMCVCLPVTVIYILMDRWILGELICKVTPFVQ 110
Db 1 LVFIISQKEMNTVNIILANLSCSDILMCVCLPVTVIYILMDRWILGELICKVTPFVQ 110
Qy 116 CFHEACVSFASVSTAINVFA---ITLDYDISVKPANILTMGRVLMISITWIFSF 172
Db 1 CFHEACVSFASVSTAINVFA---ITLDYDISVKPANILTMGRVLMISITWIFSF 172
Qy 111 C-----MSVTISFSLVIALERHQLIHPTGTPAAGSHSYLAVATVMVACF-I 159
Db 1 C-----MSVTISFSLVIALERHQLIHPTGTPAAGSHSYLAVATVMVACF-I 159
Qy 173 LPIPIEVNFFSLQSGNTWENKTL-----LCV-----STNEYYTELGMYYHLVQ 216
Db 1 LPIPIEVNFFSLQSGNTWENKTL-----LCV-----STNEYYTELGMYYHLVQ 216
Qy 160 SLPELS-----FNLINAPQISLFPNPSFDVHICMELPSENRRLAYTISLLFQYCL-- 214
Db 1 SLPELS-----FNLINAPQISLFPNPSFDVHICMELPSENRRLAYTISLLFQYCL-- 214
Qy 217 IPIFFFTVVMLITYTKILOALNIRIGTRESTGQKKKARKKTKISLTQHEATDMSQSSG 276
Db 1 IPIFFFTVVMLITYTKILOALNIRIGTRESTGQKKKARKKTKISLTQHEATDMSQSSG 276
Qy 215 -----PLLLILLCYLRIF--LRLR-----RRKDMVEQATE----- 242
Db 1 -----PLLLILLCYLRIF--LRLR-----RRKDMVEQATE----- 242
Qy 277 GRNVVGVRTSVSVIIALRRVAKRHRRERQKRVFRMSLLIISTFLCWTPTISVLNTTI 336
Db 1 GRNVVGVRTSVSVIIALRRVAKRHRRERQKRVFRMSLLIISTFLCWTPTISVLNTTI 336
Qy 243 -----ARQKARGAQ-RVNAMLVIVVAFALCWLPLNVNTIF 279
Db 1 -----ARQKARGAQ-RVNAMLVIVVAFALCWLPLNVNTIF 279
Qy 337 ---LCLGPSDLLVKRLCFLVMAYGTIFPHPLLYAFTRQKFKVLSKMKK 384
Db 1 ---LCLGPSDLLVKRLCFLVMAYGTIFPHPLLYAFTRQKFKVLSKMKK 384
Qy 280 DWYHQALPACQHDVIFSACHLT-AMASTCVNPVYVYGLNTNFQKELKATLQR 330
Db 1 DWYHQALPACQHDVIFSACHLT-AMASTCVNPVYVYGLNTNFQKELKATLQR 330

RESULT 9
ID O62059 PRELIMINARY; PRT; 643 AA.
AC O62059;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 73.0 kDa protein Y54E2A.1.
GN Y54E2A.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL032646; CAA21687.2; -.
DR EMBL; Z81465; CAA21687.2; JOINED.
DR EMBL; Z81465; CAB03866.2; -.
DR EMBL; AL032646; CAB03866.2; JOINED.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 643 AA; 72950 MW; 655E1A9B5E27927 CRC64;

Query Match 10.1%; Score 222; DB 5; Length 643;
Best Local Similarity 22.2%; Pred. No. 3.6e-11;
Matches 94; Conservative 86; Mismatches 188; Indels 56; Gaps 18;

Qy 2 CFSPI---LEINMOSESNITVRDDIDDINTNMYQPLSYPLSFQVSLTGFLMLEIVL--GL 56
Db 1 CFSPI---LEINMOSESNITVRDDIDDINTNMYQPLSYPLSFQVSLTGFLMLEIVL--GL 56
Qy 38 CTHPLRKYLEVCI---SRCTVPDD-----TVFFSMTDELFELPGLFLYTLVFLVGTI 88
Db 1 CTHPLRKYLEVCI---SRCTVPDD-----TVFFSMTDELFELPGLFLYTLVFLVGTI 88
Qy 57 GSNLTVLVLYCMKSNLINSVNIITMNLHVLDVLCVGCIPLTIVILLLSLESNTALICC 116
Db 1 GSNLTVLVLYCMKSNLINSVNIITMNLHVLDVLCVGCIPLTIVILLLSLESNTALICC 116
Qy 89 GNSWVIFVNNRFR--MRNTNIFLASLSTADLCILWFCVP-----IMFMKYSMTWSMGR 142
Db 1 GNSWVIFVNNRFR--MRNTNIFLASLSTADLCILWFCVP-----IMFMKYSMTWSMGR 142
Qy 117 FHEACVS-----FASVSTAINVFAITLDYDISVKPANRI--LTMGRVLMISITWIFS 168
Db 1 FHEACVS-----FASVSTAINVFAITLDYDISVKPANRI--LTMGRVLMISITWIFS 168
Qy 143 F--ACYSVHYIQOFCFCFSVLTMTMISPERFLAIAYPRNINWFSSIGRAKKVILLIWKSS 200
Db 1 F--ACYSVHYIQOFCFCFSVLTMTMISPERFLAIAYPRNINWFSSIGRAKKVILLIWKSS 200
Qy 169 FFSFLIPIEVNF---FSLQSGNT--WENKTLCSVTNEYYTELGMYYHLVQIPIFFFTV 224
Db 1 FFSFLIPIEVNF---FSLQSGNT--WENKTLCSVTNEYYTELGMYYHLVQIPIFFFTV 224
Qy 201 AVLAVPTAVRMDYETNLSLSQGVHWCRRRFPAQFLGYPRTSLNKAYAMYQLLLIIFPV 260
Db 1 AVLAVPTAVRMDYETNLSLSQGVHWCRRRFPAQFLGYPRTSLNKAYAMYQLLLIIFPV 260
Qy 225 VVMLITYTKILOALNIRIGTRESTGQKKKARKKTKISLTQHEATDMSOSSGRRNVVGV 284
Db 1 VVMLITYTKILOALNIRIGTRESTGQKKKARKKTKISLTQHEATDMSOSSGRRNVVGV 284
Qy 261 LTMSCYARVSAIVYKSKDRVILSQAMVAFSKAATDAVTFSGYSAIPMITTSRN----L 316
Db 1 LTMSCYARVSAIVYKSKDRVILSQAMVAFSKAATDAVTFSGYSAIPMITTSRN----L 316
Qy 285 RTSVVIILARRAVKRHRERR---BRQKRVFRMSLLIISTFLCWTPTISVLNTTILCLG- 340
Db 1 RTSVVIILARRAVKRHRERR---BRQKRVFRMSLLIISTFLCWTPTISVLNTTILCLG- 340
Qy 317 KTANNTI---KSYSHRRNNRVAEANKQIVQMLISICMYTVCWLP-TIVDELLTSFGY 371
Db 1 KTANNTI---KSYSHRRNNRVAEANKQIVQMLISICMYTVCWLP-TIVDELLTSFGY 371
Qy 341 -----PSDLLVKRLCFLVMAYGTIFPHPLLYAFTRQKFKVVK---SKMKRV--VSIVE 391
Db 1 -----PSDLLVKRLCFLVMAYGTIFPHPLLYAFTRQKFKVVK---SKMKRV--VSIVE 391
Qy 372 ICRTSNTOTLHMRMGFNALTYCQSCINPILYAFISQNFSTFTKAYSRMKSRLQGVHVS 431
Db 1 ICRTSNTOTLHMRMGFNALTYCQSCINPILYAFISQNFSTFTKAYSRMKSRLQGVHVS 431
Qy 392 ADPL 395
Db 1 ADPL 395
Qy 432 GDPL 435
Db 1 GDPL 435

RESULT 10
Qy 9297 PRELIMINARY; PRT; 422 AA.
ID Qy 9297;
AC Qy 9297;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serotonin receptor 1A.
GN HTR1A.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=GORILLA-U1;
RA Kitano T., Kobayakawa H., Saitou N.;
RT "Silver Project.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AB041405; BAA94490.1; -.

```

Q23305 Q23305 PRELIMINARY; PRT; 539 AA.
 AC Q23305;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE 2C412.1 protein.
 GN 2C412.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ainscough R.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 RL EMBL; 278067; CAB01528.1;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1.1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
 SQ SEQUENCE 539 AA; 62655 MW; F2E750CCC7FE5C2D CRC64;
 Query Match 10.2%; Score 224; DB 5; Length 539;
 Best Local Similarity 21.2%; Pred. No. 2e-11;
 Matches 89; Conservative 82; Mismatches 150; Indels 98; Gaps 14;
 QY 25 DINTNMQPLSYPLSFQVSLTGFLMLE---IVLGASNLTVLVLYCMKSNLINSVSNIT 81
 DB 11 DVNAIIQQPNDVTVLEFVRL-GYSVLYFLIIIGLVGGLIITSILMKKL--SVANIFL 67
 QY 82 MNLHVLVIVCGICPLTIVILLLESNTALICCFHEACVSVSTAINVFAITLDRI 141
 DB 68 INLAVSLLLCITAVPTTPVLAFLMKWIFGIIMCKLVPTCAQFSLVLSGLYIAIDRY 127
 QY 142 DISVKKPANRLTGMRAVLMISWIFSFELIPFIEVNFESLQSGNTWENKTL---LC- 197
 DB 128 RSIVTLPREPWSDRHAWLMTFWVAFVLAFLYISQNLKTM-----VIENVLGDFCG 183
 QY 198 -----VSTNEYTELGHYHLLVQIPIFFFTVVVMLITYTKIQLALNIRIGTRFSTG 249
 DB 184 EFNQSDSEISKLYTTSL-----LIQLII---PAIMSPCYLMILQKVQ----- 225
 QY 250 QKKARKKKTISLTTOHEATDMSQSGGRNVFGVTSVVIALLRAVKKHRRERQK 309
 DB 226 -----TDWLVDG-----SMLTAAQQA-----QTAVRKR 249
 QY 310 RVFRMSLLIISTFLCWTPIVSLN-----TTILCGPSDLLVKRLCLFLVMA 356
 DB 250 RMVYLIVLVIVMACFVPLUSANFLDLGNRFECQTVYKVLMDQMYEKL-LNVHVA 308
 QY 357 YGTFIFHPLLYAFTRQKQVLSKSM-----KKRVSVIVEADPLPNNAVIHNSWID 407
 DB 309 MTSIVNVPVLYFWMSKRRRALDDMTWLTNARRHTNVGLSFRFTSPSPSVVYRRTLE 367
 RESULT 7
 Q9N298
 ID Q9N298 PRELIMINARY; PRT; 422 AA.
 AC Q9N298;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Serotonin receptor 1A.
 GN HTR1A.

OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHIMP-220;
 RA Kitano T., Kobayakawa H., Saitou N.;
 RT "Silver Project."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AB041404; BAA94489.1;
 DR HSSP; P29274; 1MMH.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1.1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
 SQ SEQUENCE 422 AA; 46171 MW; FCD989BD0313A1A0 CRC64;
 Query Match 10.2%; Score 223; DB 6; Length 422;
 Best Local Similarity 21.8%; Pred. No. 1.8e-11;
 Matches 93; Conservative 73; Mismatches 157; Indels 104; Gaps 16;
 QY 27 NTNNMQPLSYPLSFQV---SLTQFLMLEIVLGLSGNLTVLVLYCMKSNLINSVSNITMN 83
 DB 24 NTSGISDVF--SYQVITSLLLGLTIFCAVLG---NACVVAALIALERSLQN-VANYLIGS 77
 QY 84 LHVLDVIVCGICPLTIVILLLESNTALICCFHEACVSVSTAINVFAITLDRI 143
 DB 78 LAVTDLMSVSLVLPMAALYQVLNKNLTGLQVTCDFIALDLVLCCTSSLHLCAIALDRIYA 137
 QY 144 SVKPANRI--LTGMRAVLMISWIFSFELIPFIEVNFESLQSGNTWENKTL---LLCVS 199
 DB 138 ITDPIDVVKRTPRRAALSLTLWGLFSLIPML-----GWTPEDRSPPDACTI 189
 QY 200 TNEY-----YTELGMVYHLLVQIPIFFFTVVVMLITYTKIQLALNIRIGTRFSTGQKKAR 255
 DB 190 SKDHGTYITFTGAFY-----IPL-----LLMLVLYGRIFRAARFRI-----RKTVK 231
 QY 256 KKTISLTTOHEATDMSQ-----SSGRNVFGVTSVVIALLRAVKK----- 299
 DB 232 KVETGADTRHGASPAQPKKSVNGESGRNWRGLGVSKAGGALCANGAVRQGDGALE 291
 QY 300 -----RHREREROKRVFRMSL-----RHREREROKRVFRMSL-----L 317
 DB 292 VIEVHRVGNSEKHLPLPSEAGTPCAPASFERNKERNNAEAKR--KMALAREKTVKTLGI 349
 QY 318 IISTFLCWTPIVSLNNTI-LCLGPSDLLVKRLCLFLVMA YGTFIFHPLLYAFTRQKQK 376
 DB 350 IMGTFLCWLFPFFIVALVLPFCSSCHMPTLLGAILNWLGYNSLLNPVIYAVFNKDFQN 409
 QY 377 VLKSKMK 383
 DB 410 AFKKIIR 416
 RESULT 8
 O73734
 ID O73734 PRELIMINARY; PRT; 373 AA.
 AC O73734;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE Neuropeptide Y/peptide YY receptor Yc.
 GN NPYRYC OR NPYRYC.
 OS Brachydanio rerio (zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;

| | | | |
|---|--|--|-----|
| Qy | 96 | IPLTIVILLLSLSENATLCCFHEACVSF-----ASVSTAINFAITLDRYDISVKP--A | 14 |
| Db | 108 | MPETLLENMGTFIFGTVIC---KAVSLMGVSVSVST-LNLVAIALERYSAICRPLQA | 162 |
| Qy | 149 | NRILTMGRAVLMISLWISFSEFLPIETEVNFFS-----LQSGNTWENKLLCVSTNE | 202 |
| Db | 163 | RVNQTRSHAARVILAWLLLSGL-LMVPYPVYVYVQVGPVLCQHRWPS----- | 211 |
| Qy | 203 | YYTELGMYHLLVQIPIFFFTVVVMLITYTKILQALNIRIGTRF--STGQKKKARKKTKI | 260 |
| Db | 212 | --ARVROTWSVLLMLLFFPIPGVMAVAGLLISREL--YGLRFDGDNDSDFQSRVNOG | 267 |
| Qy | 261 | SLTQHEATDMKSSSGGRNVVFCVTSVSIALLRA-----VKRHRER | 304 |
| Db | 268 | GLPGTAGPVGPHVGGCRHVTVAGEDNDGCVQLPRSLRMTLLTPTPGPGLASANOAK | 327 |
| Qy | 305 | REQKRVFRMSLLIISTFELCWTPISVLNNTTILCLGPS--DLLVKLRCLF-LVMAYGTTI | 361 |
| Db | 328 | LLAKRVVRMLLVILLFFLCWLPYISANTWRAFDGGAHRALSGAPISFIHLLSYASAC | 387 |
| Qy | 362 | FHPLLYAFTQKFEK 376 | |
| Db | 388 | VNPLVYCFMHRFRQ 402 | |
| RESULT 4 | | | |
| Qy | 057463 | PRELIMINARY; PRT; 375 AA. | |
| Id | 057463 | | |
| AC | 057463; | | |
| DT | 01-JUN-1998 (T-EMBLrel. 06, Created) | | |
| DT | 01-JUN-1998 (T-EMBLrel. 06, Last sequence update) | | |
| DT | 01-DEC-2001 (T-EMBLrel. 19, Last annotation update) | | |
| DE | Neuropeptide Y /peptide YY receptor Yb. | | |
| GN | NPYRYb. | | |
| OS | Brachydanio rerio (Zebrafish) (Zebra danio). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; | | |
| OC | Cyprinidae; Danio. | | |
| OX | NCBI_TaxId=7955; | | |
| RP | [1] | | |
| RN | SEQUENCE FROM N.A. | | |
| RA | Lundell I., Berglund M.M., Starback P., Salaneck S., Gehlert D.R., | | |
| RA | Larhammar D.; | | |
| RT | "Cloning and characterization of a novel neuropeptide Y (NPY) receptor | | |
| RT | subtype in the zebrafish."; | | |
| RL | DNA Cell Biol. 0:0-0(1997). | | |
| DR | EMBL; AF030245; AAB94616.1; -. | | |
| DR | ZFIN; ZDB-GENE-980526-208; npyryb. | | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | | |
| DR | Pfam; PF00001; 7tm_1; 1. | | |
| DR | PRINTS; PR00237; GPCRHHODPSN. | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECEP_F1.1; UNKNOWN_1. | | |
| DR | PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1. | | |
| KW | Receptor. | | |
| SQ | SEQUENCE 375 AA; 42116 MW; 4B0C9551131A14D1 CRC64; | | |
| Query Match 10.5%; Score 229.5; DB 13; Length 375; | | | |
| Best Local Similarity 21.5%; Pred. No. 4.4e-12; | | | |
| Matches 81; Conservative 76; Mismatches 114; Indels 105; Gaps | | | |
| Qy | 42 | VSLTGFLMLE-----IVILGSGNLTVLVLYCMKSNLINSYNIITMLNHLVDVLCVGPIC | 97 |
| Db | 22 | LSSTFLIVAYSTMALVGLVGN-TCLVWVITRQKEMRNVTNFIYNLSCDILVCLVCLIP | 80 |
| Qy | 98 | LTIVILL-----SLESNTALICCFHEACVSFASVSTAINVFA---ITLDRYDISVKP | 147 |
| Db | 81 | VTIITYTLMDRWILGEALCKVTPFVQC-----MSVTVSIFSMVLIALERHQIILHP | 130 |
| Qy | 148 | ANRLTMGRAVLMISLWISFSEFLPIETEVNFFSILQSGNTWENKTL-----LCV | 198 |
| Db | 131 | TGWKPVVRHSTVLAVVLIWITACE-LSLPSF---FNLINSPFNLSLDFNPFSDHFIC | 186 |
| Qy | 199 | STNEYYTELG--MYHLLVQIPIFFFTVVVMLITYTKILQALNIRIGTRFSGCKKARK | 256 |

RESULT 6

```

Db 116 TTTTYPGTSVSVSTNLVAISLERYGALCPQLSRVWTKSHALKVIAATWCLSTIMTP 175
Qy 170 ---FSLPIPFIEVN-----FFSLQSG---NTWENKTLVCVSTNEYTELGMYYHLLVQ 216
Db 176 YPIYSLNLPVFTKNNQNTANMCRFLLPSPDAQMSQWTFLLLI----- 216
Qy 217 IPIFFTVVVMILITVKILQALNIRIGTRFSTGQKKKARKKTI-----SLTQHEA 268
Db 217 ---LFLIPGVWVAYGLI--SLEYQGIKFDASQKSAKRLSSGGGGSSSRVED 272
Qy 269 TD-----MSQSSGGGRNVFVGVTSV--IIALRAVKKHRRERQ 308
Db 273 SDGCVLQKSRPRKLEQLQSLTSSSGR--INRISSGSAANLIA-----K 316
Qy 309 KRVRMSLLIISTFLCWTPTISVLNT-----TI-----LCLGPSDLLVKLRCLFVLMAYG 358
Db 317 KRVRMLIVIVVFLCWPFISSANAWRAYDTVSAEKHLSTGTPISFI-----LLLSYT 369
Qy 359 TTIFHPLIYAFTROKFQ 375
Db 370 SSCVNPIIYCFMKNRFR 386

RESULT 2
Q9DBV6 PRELIMINARY; PRT; 436 AA.
AC Q9DBV6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Adult male lung cDNA, RIKEN full-length enriched library,
DE clone:1200012013, full insert sequence.
GN CCKAR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001)
DR EMBL; AK004730; BAB23512.1; -
DR HSSP; P02699; IF88.
DR MGI; MGI:99478; Cckar.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01570; NFFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
SQ SEQUENCE 436 AA; 48446 MW; 49B4AD57F080F08A CRC64;

Query Match 12.1%; Score 266; DB 11; Length 436;
Best Local Similarity 23.8%; Pred. No. 3.2e-15;
Matches 104; Conservative 85; Mismatches 128; Indels 120; Gaps 20;

Qy 6 ILEINMGSESNITVRDDIDDTNNMV---QP---LSYPLSFQSVSLTGFLMLEIVLGLGSN 59
Db 3 VVDSLMLNGSNITPPCELGLENETLFCLDQPHPSKEWQSAVOILLYSFIFLUSVLGDTLV 62
Qy 60 LTVLVLYCMKSNLINSVSNIIIMNLHVLVDIICVCIPLTIVILLLESNTALICCFHE 119
Db 63 ITVLI---RNKMRVTNIFLLSLAVSDMLCLFCMPFNLPNLLKDFIFGSVC---K 115
Qy 120 ACVSFASVSTAINVF---AITLDRYDISVKP--ANRI--LTMGRAVMLMISINIFSP---- 169
Db 116 TTTTYPGTSVSVSTNLVAISLERYGALCPQLSRVWTKSHALKVIAATWCLSTIMTP 175
Qy 170 ---FSLPIPFIEVN-----FFSLQSG---NTWENKTLVCVSTNEYTELGMYYHLLVQ 216
Db 176 YPIYSLNLPVFTKNNQNTANMCRFLLPSPDAQMSQWTFLLLI----- 216
Qy 217 IPIFFTVVVMILITVKILQALNIRIGTRFSTGQKKKARKKTI-----SLTQHEA 268
Db 217 ---LFLIPGVWVAYGLI--SLEYQGIKFDASQKSAKRLSSGGGGSSSRVED 272
Qy 269 TD-----MSQSSGGGRNVFVGVTSV--IIALRAVKKHRRERQ 308
Db 273 SDGCVLQKSRPRKLEQLQSLTSSSGR--INRISSGSAANLIA-----K 316
Qy 309 KRVRMSLLIISTFLCWTPTISVLNT-----TI-----LCLGPSDLLVKLRCLFVLMAYG 358
Db 317 KRVRMLIVIVVFLCWPFISSANAWRAYDTVSAEKHLSTGTPISFI-----LLLSYT 369
Qy 359 TTIFHPLIYAFTROKFQ 375
Db 370 SSCVNPIIYCFMKNRFR 386

RESULT 3
P89005 PRELIMINARY; PRT; 450 AA.
AC P89005;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Gastrin/cholecystokinin-B receptor (fragment).
DE Protons naltalensis (African soft-furred rat) (Mastomys natalensis).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
DE Mastomys
OX NCBI_TaxID=10112;
RN [1]
RP SEQUENCE FROM N.A.
RA Luque E.A., Tang L.H., Modlin I.M.;
RA "Gastrin receptor expression in Mastomys natalensis.";
RT Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49834; AAB4677.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 450 AA; 48792 MW; 333F4C368B7A0A97 CRC64;

Query Match 10.7%; Score 235.5; DB 11; Length 450;
Best Local Similarity 23.5%; Pred. No. 1.6e-12;
Matches 88; Conservative 71; Mismatches 149; Indels 67; Gaps 13;

Qy 36 YPLSFQSVSLTGFLMLEIVLGLGSNLTVLVLCMKSNLINSVNIITMNLHVLDIIVCGC 95
Db 61 YAVIFLMSIGGNMLIIVLGL-----SRLRTVTNNAFLSLAVSLLAVAC 107
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:59:53 ; Search time 35 Seconds
(without alignments)
2549.097 Million cell updates/sec

Title: US-09-845-721-2
Perfect score: 2192
Sequence: 1 MCFSPLEINMQSESNITVR.....ITFDESEIRRLVPQVYTD 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 267 | 12.2 | 436 | 11 | Q8VCC7 | Q8vcc7 mus musculus |
| 2 | 266 | 12.1 | 436 | 11 | Q8DBV6 | Q8dbv6 mus musculus |
| 3 | 235.5 | 10.7 | 450 | 11 | P89005 | P89005 proomys nat |
| 4 | 229.5 | 10.5 | 375 | 13 | O57463 | O57463 brachydanio |
| 5 | 225 | 10.3 | 377 | 13 | Q8QGM3 | Q8qgm3 gallus gall |
| 6 | 224 | 10.2 | 539 | 5 | Q23305 | Q23305 caenorhabdi |
| 7 | 223 | 10.2 | 422 | 6 | Q9N298 | Q9n298 pan troglod |
| 8 | 222 | 10.1 | 373 | 13 | Q73734 | Q73734 brachydanio |
| 9 | 222 | 10.1 | 643 | 5 | O62059 | O62059 caenorhabdi |
| 10 | 221 | 10.1 | 422 | 6 | Q9N297 | Q9n297 gorilla gor |
| 11 | 218 | 9.9 | 396 | 4 | Q92492 | Q92492 homo sapien |
| 12 | 217 | 9.9 | 422 | 6 | Q9N296 | Q9n296 pongo pygma |
| 13 | 217 | 9.9 | 447 | 4 | Q16144 | Q16144 homo sapien |
| 14 | 217 | 9.9 | 458 | 13 | Q98843 | Q98843 anguilla an |
| 15 | 216 | 9.9 | 399 | 5 | Q9NG02 | Q9ng02 apis mellif |
| 16 | 215.5 | 9.8 | 518 | 6 | Q9MYI8 | Q9myi8 oryctolagus |

| | | | | | | |
|----|-------|-----|-----|----|--------|---------------------|
| 17 | 213.5 | 9.7 | 464 | 5 | O9G054 | O9g054 aedes aegyp |
| 18 | 213 | 9.7 | 396 | 4 | O75824 | O75824 homo sapien |
| 19 | 211.5 | 9.6 | 390 | 6 | Q8WNV9 | Q8wnv9 sus scrofa |
| 20 | 211.5 | 9.6 | 440 | 5 | Q9N324 | Q9n324 caenorhabdi |
| 21 | 211 | 9.6 | 377 | 5 | O02464 | O02464 manduca sex |
| 22 | 209.5 | 9.6 | 365 | 13 | Q8UVW7 | Q8uvw7 lampetra fl |
| 23 | 209 | 9.5 | 515 | 11 | O9DBL0 | O9dbl0 mus musculu |
| 24 | 208.5 | 9.5 | 516 | 4 | Q9NYK7 | Q9nyk7 homo sapien |
| 25 | 208 | 9.5 | 379 | 5 | O76124 | O76124 papilio xut |
| 26 | 206 | 9.4 | 380 | 5 | Q9UAM7 | Q9uam7 papilio gla |
| 27 | 205.5 | 9.4 | 374 | 13 | Q9VHX1 | Q9vhx1 gadus morhua |
| 28 | 205.5 | 9.4 | 402 | 5 | Q964E5 | Q964e5 dugesia tig |
| 29 | 205.5 | 9.4 | 516 | 4 | Q96LC6 | Q96lc6 homo sapien |
| 30 | 205 | 9.4 | 408 | 13 | Q98998 | Q98998 xenopus lae |
| 31 | 204 | 9.3 | 383 | 6 | Q9GK75 | Q9gk75 macaca mula |
| 32 | 203 | 9.3 | 487 | 6 | Q9N2B1 | Q9n2b1 gorilla gor |
| 33 | 201.5 | 9.2 | 310 | 13 | Q8QGK6 | Q8qgx6 fuqu rubrip |
| 34 | 201.5 | 9.2 | 446 | 6 | Q8WND7 | Q8wnd7 bos taurus |
| 35 | 201 | 9.2 | 474 | 11 | Q8R456 | Q8r456 rattus norv |
| 36 | 200 | 9.1 | 377 | 13 | O73733 | O73733 brachydanio |
| 37 | 200 | 9.1 | 477 | 5 | Q9BMA9 | Q9bma9 mamestra br |
| 38 | 199 | 9.1 | 419 | 5 | O77254 | O77254 boophilus m |
| 39 | 199 | 9.1 | 457 | 5 | Q18534 | Q18534 caenorhabdi |
| 40 | 197.5 | 9.0 | 390 | 4 | Q96LD9 | Q96ld9 homo sapien |
| 41 | 197 | 9.0 | 372 | 13 | P79945 | P79945 xenopus lae |
| 42 | 197 | 9.0 | 471 | 13 | Q8UUG8 | Q8uug8 tetraodon f |
| 43 | 197 | 9.0 | 487 | 6 | Q9N2B2 | Q9n2b2 pan troglod |
| 44 | 196 | 8.9 | 381 | 5 | Q95YI3 | Q95yi3 bombyx mori |
| 45 | 196 | 8.9 | 660 | 5 | Q9VFW5 | Q9vfw5 drosophila |

ALIGNMENTS

RESULT 1

| | | | | |
|--------|--|--------------|------|---------|
| Q8VCC7 | Q8VCC7 | PRELIMINARY; | PRT; | 436 AA. |
| AC | Q8VCC7; | | | |
| DT | 01-MAR-2002 (TReMBLrel. 20, Created) | | | |
| DT | 01-MAR-2002 (TReMBLrel. 20, Last sequence update) | | | |
| DT | 01-JUN-2002 (TReMBLrel. 21, Last annotation update) | | | |
| DE | Cholecystokinin A receptor. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=KIDNEY; | | | |
| RA | Strausberg R.; | | | |
| RL | Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; BC020534; AAH20534.1; - | | | |
| DR | InterPro; IPR000276; GPCR_Rhodopsn. | | | |
| DR | Pfam; PF00001; 7tm1.1; 1. | | | |
| DR | PRINTS; PRO0237; GPCR_RHODOPSN. | | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1. | | | |
| DR | PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1. | | | |
| KW | Receptor. | | | |
| SQ | SEQUENCE 436 AA; 48396 MW; FD35D8453B792F79 CRC64; | | | |

Query Match 12.2%; Score 267; DB 11; Length 436;
Best Local Similarity 23.8%; Pred. No. 2.6e-15;
Matches 104; Conservative 85; Mismatches 128; Indels 120; Gaps 20;

| | | | |
|----|-----|--|-----|
| Qy | 6 | ILEINMQSESNITVRDIDDINTNMV-----QP-LSYPLSFQVSLTGLMLEIVLGLGSN | 59 |
| Db | 3 | VVDSLLMNGSNITPPCELGLENETFLCDLPQPSKEQSAVQILLYSFILSLVIGNTLV | 62 |
| Qy | 60 | LTVLVLCMKSNLINSVNIITMLHVLDTICVGCPLTIVILLLSLESNTALICFHE | 119 |
| Db | 63 | ITVLV-----RNKRMTVTNIFLLSLAVSDMLCLFCFPNLPNLKDFIGSAVC---K | 115 |
| Qy | 120 | ACVSFASVSTAINVF---ATLDRYDISVRP-ANRI-LTMGRVLMISIWFSF----- | 169 |

This Page Blank (uspto)

```
FT TRANSMEM 132 150 3 (POTENTIAL).
FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 189 4 (POTENTIAL).
FT DOMAIN 190 219 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 220 242 5 (POTENTIAL).
FT DOMAIN 243 333 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 334 355 6 (POTENTIAL).
FT DOMAIN 356 373 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 374 394 7 (POTENTIAL).
FT DOMAIN 395 447 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 127 205 BY SIMILARITY.
FT LIPID 408 408 PALMITATE (BY SIMILARITY).
FT CONFLICT 171 171 A -> P (IN REF. 5).
FT CONFLICT 249 249 L -> V (IN REF. 5).
SQ SEQUENCE 447 AA; 48419 MW; BAEFAD4C1F85915 CRC64;

Query Match 9.9%; Score 218; DB 1; Length 447;
Best Local Similarity 22.8%; Pred. No. 6.1e-08;
Matches 88; Conservative 70; Mismatches 136; Indels 92; Gaps 15;

QY 36 YPLSFQVSLTGFLMLBEIVLGLGSLNLTFLVLYCMKSLNLSVSNITMTNHLHVLVDVIVCGC 95
DB 61 YAVIFLMSVGGNMLIIVLGL-----SRRLRTVTFNAFLSLAVSDELLAVAC 107

QY 96 IPTIVILLLESNTALICCFHEACVSF-----ASVSTAINVFAITLDRYDISVRP--A 148
DB 108 MPFTLLPNLMGTIFGTVIC---KAVSYLMGVSVSVST-LSLVAIALERYSAICRPLQA 162

QY 149 NRLLTNGRAVLMISWIFSPFLPFIEVNEFS-----LQSGNTWENKTLICVSTNE 202
DB 163 RVWQTRSHAARVIVATWLLSGL-LMVPYPVYTVVQVGPVQLQCVHRWPS----- 211

QY 203 YYTELGMYYHLVQIPFEFTVVVMLTITKILQALNIRIGTRF-----STGQKKKARK 257
DB 212 --ARVQOTWSVLLLLLFIPGVVMAVAYGLISREL--YLGFRFDGSDSDSOSVRNQG 267

QY 258 KTISLATQH-----EATDMSQSSGGRNVVF-----GVRTSVSVIIA 293
DB 268 GLPGAVHQNGRCRPEYGAIVEDSDGCVQLPRSRPALELTALTAPGPGSRPTQAKLLA 327

QY 294 LRAVKRRHRRERQKRVFRMSLLIISTELLCTWTPISVLTTLICLGPS--DLLVKLRLC 351
DB 328 -----KKRVVRLVIVVLFELCWLPLVYSANTWRAFDGPGAHRLSGAPIS 373

QY 352 FL-VMAYGTIFHPHLLYAFTRQEQK 376
DB 374 FIHLLSYACVNPLVYCFMHRFRQ 399
```

Search completed: December 4, 2002, 16:01:36
Job time : 15 secs

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

| | | | |
|----|--|-----------------------------------|--|
| | EMBL; M28269; AAA36440.1; -; | | |
| DR | EMBL; X13556; CAA31908.1; -; | | |
| DR | EMBL; X57829; CAA40962.1; -; | | |
| DR | EMBL; M83181; AAA66493.1; -; | | |
| DR | PIR; S07343; S07343. | | |
| DR | PIR; S31438; S31438. | | |
| DR | HSP; P29274; IMMH. | | |
| DR | Genew; HGNC:5286; HTRIA. | | |
| DR | MIM; 109760; -; | | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | | |
| DR | Pfam; PF00001; 7tm_1; 1. | | |
| DR | PRINTS; PR00237; GPCRRHODPSN. | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1. | | |
| DR | PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1. | | |
| KW | G-protein coupled receptor; Transmembrane; Glycoprotein; | | |
| KW | Multigene family; Polymorphism. | | |
| FT | DOMAIN 1 36 | EXTRACELLULAR (POTENTIAL). | |
| FT | TRANSMEM 37 62 | 1 (POTENTIAL). | |
| FT | DOMAIN 63 73 | CYTOPLASMIC (POTENTIAL). | |
| FT | TRANSMEM 74 98 | 2 (POTENTIAL). | |
| FT | DOMAIN 99 109 | EXTRACELLULAR (POTENTIAL). | |
| FT | TRANSMEM 110 132 | 3 (POTENTIAL). | |
| FT | DOMAIN 133 152 | CYTOPLASMIC (POTENTIAL). | |
| FT | TRANSMEM 153 178 | 4 (POTENTIAL). | |
| FT | DOMAIN 179 191 | EXTRACELLULAR (POTENTIAL). | |
| FT | TRANSMEM 192 217 | 5 (POTENTIAL). | |
| FT | DOMAIN 218 345 | CYTOPLASMIC (POTENTIAL). | |
| FT | TRANSMEM 346 367 | 6 (POTENTIAL). | |
| FT | DOMAIN 368 378 | EXTRACELLULAR (POTENTIAL). | |
| FT | TRANSMEM 379 403 | 7 (POTENTIAL). | |
| FT | DOMAIN 404 422 | CYTOPLASMIC (POTENTIAL). | |
| FT | CARBOHYD 10 10 | N-LINKED (GLCNAC. .) (POTENTIAL). | |
| FT | CARBOHYD 11 11 | N-LINKED (GLCNAC. .) (POTENTIAL). | |
| FT | CARBOHYD 24 24 | N-LINKED (GLCNAC. .) (POTENTIAL). | |
| FT | DISULFID 109 187 | BY SIMILARITY. | |
| FT | VARIANT 16 16 | P -> L. | |
| FT | VARIANT 22 22 | /FTid=VAR_003446. | |
| FT | VARIANT 28 28 | G -> S (IN DBSNP:1799920). | |
| FT | VARIANT 184 184 | /FTid=VAR_011826. | |
| FT | VARIANT 220 220 | I -> V (IN DBSNP:1799921). | |
| FT | VARIANT 272 272 | /FTid=VAR_011827. | |
| FT | VARIANT 273 273 | P -> L (IN DBSNP:1800043). | |
| FT | VARIANT 152 154 | /FTid=VAR_011828. | |
| FT | VARIANT 172 172 | R -> L (IN DBSNP:1800044). | |
| FT | VARIANT 418 418 | /FTid=VAR_011829. | |
| FT | VARIANT 422 AA; 46106 MW; 762664FC6CFDBF CRC64; | G -> D. | |
| FT | VARIANT 152 154 | /FTid=VAR_003447. | |
| FT | VARIANT 172 172 | G -> D (IN DBSNP:1800042). | |
| FT | VARIANT 418 418 | /FTid=VAR_011830. | |
| FT | VARIANT 422 AA; 46106 MW; 762664FC6CFDBF CRC64; | RAA -> PR (IN REF. 1). | |
| FT | VARIANT 152 154 | M -> I (IN REF. 1). | |
| FT | VARIANT 172 172 | K -> N (IN REF. 1). | |
| FT | VARIANT 418 418 | | |
| FT | VARIANT 422 AA; 46106 MW; 762664FC6CFDBF CRC64; | | |

```

Query Match      10.1%; Score 222; DB 1; Length 422;
Best Local Similarity 21.9%; Pred. No. 3.1e-08;
Matches 91; Conservative 70; Mismatches 153; Indels 102; Gaps 15;

Qy 38 LSPQV---STGTFLMLEIVLGLGSLNLVLYVYCKMSKLNLSVSIITMNNHLVLDVITCVG 94
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 33 VSYQVITSLLTGTLIFCAVLG---NACVWAIAALERSLQN-VANYLIGSLAVTDLMYSVL 88

Qy 95 CIPLTIVILLLSLSNTALICCFHCACVSPASVSTAINVFAITLDLDYDI SVKPPANRI--L 152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 VLPMAALYQVNLKWTGLQVTCDLFTALDVLDCCTSSIIHLCAIDRLAWYATDPTDYYNKR 148

```

| | | | |
|----|-----|---|-----|
| Qy | 153 | TMGRAVLMISIWIFSFSLPIPIEYFNFFSLOGNTWENKT--LILCVSTNEY-----YTE | 206 |
| Dd | 153 | : : : : : : : : : : : : : | |
| Db | 149 | TPRRAALISLTWLIGFLISPPML-----GWRTPEDRSOPDACTISKDHGYTIYST | 200 |
| Qy | 207 | LGMYYHLLVIPFIFFVVVMVLITYTKILOALNIRIGTRFSTCQKKARKKKTISLTTOH | 266 |
| Dd | 207 | : : : : : : : : : : : : : | |
| Db | 201 | FGAFY-----IPL-----LMLLVLYGRIFAARFRI-----RKTVKKVEKTGDTRH | 242 |
| Qy | 267 | EATOWSQ-----SSGGRNVVFVRTSVSVIIALRAVK----- | 299 |
| Dd | 243 | GASPAPOKKSVNGESSRWRLGVESKAGCALCANAVRGDDGALEVIEWHRVGN SK | 302 |
| Qy | 300 | -----RHRRERRQRKVFRMSL-----LIISTFLLCWTP | 328 |
| Dd | 303 | EHLPLPSEAGPTPCAPASFERKNERNAAKR--KMALARERKVTKTLGIIMGTFILCWLP | 360 |
| Qy | 329 | ISVLNTWI-LCLGPSDLLVKLRCLFLVMAYGTTFIFHPLLYAFTRQKFQVLAKMK | 393 |
| Dd | 361 | FFIVALLVPFCRESSCHMPTLLGAILNWGLYSNLLNPVIYVFNKDFONAKFKIK | 416 |
| | | ::: ::: : : : : : : : : : : : : : : : | |
| | | RESULT 14 | |
| | | SH5B_MOUSE STANDARD; PRT; 370 AA. | |
| | | ID AC P31387; | |
| | | DT 01-JUL-1993 (Rel. 26, Created) | |
| | | DT 01-JUL-1993 (Rel. 26, Last sequence update) | |
| | | DT 01-NOV-1995 (Rel. 32, Last annotation update) | |
| | | DE 5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor). | |
| | | GN HTR5B OR 5HT5B. | |
| | | OS Mus musculus (Mouse). | |
| | | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | | OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| | | NCBI_TaxId=10090; | |
| | | EN [1] | |
| | | RP SEQUENCE FROM N.A. | |
| | | KC TISSUE=Brain; | |
| | | RX MEDLINE=93196607; PubMed=8450829; | |
| | | RA Matthes H., Boschert U., Amlaiky N., Grailhe R., Plassat J.-L., | |
| | | RA Muscatelli F., Mattei M.-G., Hen R.; | |
| | | RT "Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors | |
| | | define a new family of serotonin receptors: cloning, functional | |
| | | expression, and chromosomal localization."; | |
| | | RL Mol. Pharmacol. 43:313-319(1993). | |
| | | CC | |
| | | -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR | |
| | | 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION | |
| | | AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF | |
| | | THIS RECEPTOR IS MEDIATED BY G PROTEINS. PROBABLY INVOLVED IN | |
| | | ANXIETY AND DEPRESSION. | |
| | | CC | |
| | | -!- SUBCELLULAR LOCATION: Integral membrane protein.. | |
| | | -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE CENTRAL NERVOU | |
| | | SYSTEM; IN THE HIPPOCAMPUS, HABENULA, AND THE DORAL RAPHE. | |
| | | CC | |
| | | -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | |
| | | STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS. | |
| | | CC | |
| | | THIS SWISS-PROT entry is copyright. It is produced through a collabor | |
| | | between the Swiss Institute of Bioinformatics and the EMBL Outstat | |
| | | the European Bioinformatics Institute. There are no restrictions on | |
| | | use by non-profit institutions as long as its content is in no | |
| | | modified and this statement is not removed. Usage by and for comm | |
| | | entities requires a license agreement (See http://www.isb-sib.ch/anno | |
| | | or send an email to license@isb-sib.ch). | |
| | | CC | |
| | | EMBL; X69867; CAA49501.1; - | |
| | | DR MGD; MG1:96284; Htr5b. | |
| | | DR InterPro: IPR000276; GPCR_Rhodpsn. | |
| | | DR Pfam: PF000001; 7tm.1; 1. | |
| | | DR PRINTS; PR00237; GPCRHHODPSN. | |
| | | DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1. | |
| | | DR PROSITE; PSS0262; G_PROTEIN_RECEP_FL_2; 1. | |
| | | DR KWP Multigene family. Transmembrane; Glycoprotein; | |

DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family.


```

QY 36 YPLSFQVSLTGLFMLEIVLGLSGLNLTIVLVLYCMKSNLINSVNIITMNLHVLVDVLCVGC 95
Db 61 YAVIFLMSGGNNLITVLGL-----SRLRTVTNAFLLSLAVSDLLAVAC 107
QY 96 IPLTIVILLLESNTALCCFHEACVSP-----ASVSTAINVFAITLDRIYDISVKP--A 148
Db 108 MFTLPLNLMGTFIFGTVIC-----KAVSYLMGVSVSVST-LNLVALALERYSAICRPLQA 162
QY 149 NRILTMGRAVLMMSIWIFSEFPIEFVNFES-----LQSGNTWENKTLILCVSTNE 202
Db 163 RVWQTSRSHARVILATWLLSGL-LMVPYPTVTVQVGPVRLQCMHRWFS----- 211
QY 203 YTELOMGYHLLVQIPIFFTVVVMILITYTKILQALNIRIGTRF--STGQKKKARKKTTI 260
Db 212 --ARVQTSWLLMLLFFIPGVMAVAYGLISREL--YLGRLFDGDNDSQTSVRNQG 267
QY 261 SLTQHEATDMSSGGRNVFVGRTSVSVIATLRA-----VKRHRER 304
Db 268 GLPGGTAPGVPQHGCRHVTVAGEDNOCGYVOLPKSRLEMTLTPTTGPGLASANOAK 327
QY 305 RERQKRVFRMSLLIISTFLLCWTPISVLNLTILCLGPS--DLLVKRLCLFL-VMAYGTTI 361
Db 328 LLAKKRVVRMLLVILLFFLCWLPISYANTWCAFDGPGAHRAISGAPISFIHLLSYASAC 387
QY 362 FPHLLYAFTRQFQK 376
Db 388 VNPLVYCFMHRFRQ 402

RESULT 11
GASR_MOUSE STANDARD; PRT; 453 AA.
AC P56481;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gastrin/cholecystokinin type B receptor (CKK-B receptor) (CKK-BR).
GN CCKBR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Koplin A.S.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20318371; PubMed=10860839;
RA Lay J.M., Jenkins C., Frlis-Hansen L., Samuelson L.C.;
RT "Structure and developmental expression of the mouse CCK-B receptor gene.";
RL Biochem. Biophys. Res. Commun. 272:837-842(2000).
CC -!- FUNCTION: RECEPTOR FOR GASTRIN AND CHOLECYSTOKININ. THE CCK-B
CC RECEPTORS OCCUR THROUGHOUT THE CENTRAL NERVOUS SYSTEM WHERE THEY
CC MODULATE ANXIETY, ANALGESIA, AROUSAL, AND NEUROLEPTIC ACTIVITY.
CC THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS
CC THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF019371; AAB71863.1; -.
DR EMBL; AF264178; AAG09801.1; -.

```

```

DR EMBL; AF264177; AAG09801.1; JOINED.
DR MGD; MGI:99479; Cckbr.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS02662; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 57 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 58 79 1 (POTENTIAL).
FT DOMAIN 80 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 131 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 132 150 3 (POTENTIAL).
FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 189 4 (POTENTIAL).
FT DOMAIN 190 219 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 220 242 5 (POTENTIAL).
FT DOMAIN 243 339 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 340 361 6 (POTENTIAL).
FT DOMAIN 362 379 7 (POTENTIAL).
FT TRANSMEM 380 400 7 (POTENTIAL).
FT DOMAIN 401 453 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 127 205 N-LINKED (GLCNAC. .) (POTENTIAL).
FT LIPID 414 414 BY SIMILARITY.
FT SEQUENCE 453 AA; 49171 MW; DFD21432C33405C CRC64;

Query Match 10.3%; Score 226; DB 1; Length 453;
Best Local Similarity 22.6%; Pred. No. 1.8e-08;
Matches 93; Conservative 78; Mismatches 159; Indels 82; Gaps 15;

QY 2 CESPILQINMQSESNTVDDDDIDINTNMVQPLSYPLSFQVSLTGLFMLEIVLGLSGLT 61
Db 39 CETPRIRGTGTRLELTIRTL-----YAVIFLMSVGGNVLIIIVLGL----- 81
QY 62 VLVLYCMKSNLINSVNIITMNLHVLVDVLCVGCIPLTIVILLLESNTALCCFHEAC 121
Db 82 -----SRLRTVTNAFLLSLAVSDLLAVACMFTLTLPNLMGTFIFGTVIC-----KA 129
QY 122 VSF-----ASVSTAINVFAITLDRIYDISVKP--ANRILTMGRAVLMMSIWIFSEFSLI 174
Db 130 VSYLMGVSVSVST-LNLVALALERYSAICRPLQARVWQTSRSHARVILATWLLSGL-LMW 187
QY 175 PTIEVNFES-----LQSGNTWENKTLILCVSTNEYTELMGYHLLVQIPIFFTVVVM 228
Db 188 PVPVTVTVQVGPVRLQCMHLPSE-----RVQQMMSVLLILLFFIPGVVMA 235
QY 229 ITVTKILQALNIRIGTRF--STGQKKKARKKTTISLTQHEATDMSSGGRNVFVG-- 284
Db 236 VAYGLISREL--YLGRLFDGDNDSQTSVRNNOGGLPGGAAAPGVHQNGGCRHVTSLTG 293
QY 285 RTSVSVIIILRA-----VKRHRERERQKRVFRMSLLIISTFLLCWT 327
Db 294 EDSGCVQVQLPSRLEMTLTPTTGPGLPRPNQAKLLAKRVVRMLLVILLFFVCWL 353
QY 328 PISVLNLTILCLGPS--DLLVKRLCLFL-VMAYGTTIFPHPLLYAFTRQFQK 376
Db 354 PVSANTWRAFDGPGARRALAGAPISFIHLLSYTSACANPLVYCFMHRFRQ 405

RESULT 12
GASR_RAT
ID GASR_RAT STANDARD; PRT; 452 AA.
AC P30553;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gastrin/cholecystokinin type B receptor (CCK-B receptor) (CCK-BR).
GN CCKBR.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D86521; BAA13104.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECF_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 35 57 1 (POTENTIAL).
FT DOMAIN 58 69 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 70 90 2 (POTENTIAL).
FT DOMAIN 91 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 132 3 (POTENTIAL).
FT DOMAIN 133 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 174 4 (POTENTIAL).
FT DOMAIN 175 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 230 5 (POTENTIAL).
FT DOMAIN 231 261 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 262 284 6 (POTENTIAL).
FT DOMAIN 285 297 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 298 321 7 (POTENTIAL).
FT DOMAIN 322 371 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 109 196 BY SIMILARITY.
FT LIPID 336 336 PALMITATE (POTENTIAL).
SQ SEQUENCE 371 AA; 42447 MW; 5847E358E320F24 CRC64;

Query Match 10.8%; Score 237.5; DB 1; Length 371;
Best Local Similarity 23.0%; Pred. No. 2.5e-09;
Matches 87; Conservative 59; Mismatches 128; Indels 105; Gaps 15;

OY 52 IVLGSGSLTVLVLCMKSNLSNSVSIITMNLHVDVICGICPLTVLLLS--LES 109
DB 46 LIMGICGNLSLTIIFKQREAGQNTNLIANLSLSDILVCVWCIPFPIYIIMDRWIFG 105
OY 110 NTALCCPHEACVSPASVSTAI-NVFATLDRIYDSVSPANNILPMGRAVLMISIWIFS 168
DB 106 NT---MCKLTSYQVSISVSISFSLVLAIERYQLIVNPRGKPSASHAYGIMLIWLFS 162
OY 169 FFSFLIPFI-----EVNFESLSQSGNTWENKTLGV-----STNE--YYTELGYHYHLV 215
DB 163 LL-LSPIPLLVLHLDPEPNLSLPTDLYSHVVCVVEHWPSKTNQLLYSTSLIMLYP-- 219
OY 216 QIPIFFTVVVMILTYIKILQALNIRIGTRFSTGOKKARKKKTISLTQHEATDMSQSS 275
DB 220 -VPLGF-----MFICYLKIIVCLHR----- 239
OY 276 GGRNVVFGVTSVSVIALLRAVKKRRRERRER---QKRVFRMSLLIISTFLCWTPISVL 332
DB 240 -----NSKIDRRRENSRLTENKRNITMLISIVVTFACWLPPLNTE 280
OY 333 NT-----TILCLGSDLLVLKRLCLFLVAYGCTTFPHLLYAFTRQKQK----- 376
DB 281 NVIFDWMYHEVLMSCHHDLV--FAICHLV--AMVSTCINPLFYGLFNRNFQKDLVLIHCL 337
OY 377 --VLKSKMKRNVSVIAD 393
DB 338 CPALRERYENIAISLHTD 356

RESULT 10
GASR_PRANA

ID AC GASR_PRANA STANDARD; PRT; 450 AA.
DT P30796;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gasrin/cholecystokinin type B receptor (CKK-B receptor) (CKK-BR).
GN CCKBR.
OS *Prionys natalensis* (African soft-furred rat) (*Mastomys natalensis*).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mastomys.
OX NCBI_TaxID=10112;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92411082; PubMed=1530611;
RA Nakata H., Matsui T., Ito M., Taniguchi T., Naribayashi Y., Arima N.,
RA Nakamura A., Kinoshita Y., Chihara K., Hosoda S., Chiba T.;
RT "Cloning and characterization of gastrin receptor from ECL carcinoid
RL tumor of *Mastomys natalensis*.";
CC Biochem. Biophys. Res. Commun. 187:1151-1157(1992).
CC -!- FUNCTION: RECEPTOR FOR GASTRIN AND CHOLECYSTOKININ. THE CKK-B
CC RECEPTORS OCCUR THROUGHOUT THE CENTRAL NERVOUS SYSTEM WHERE THEY
CC MODULATE ANXIETY, ANALGESIA, AROUSAL, AND NEUROLEPTIC ACTIVITY.
CC THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS
CC THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STOMACH AND BRAIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: D12817; BAA02250.1; -
PIR: JQ1614; JQ1614.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN
DR PROSITE: PS00237; G-PROTEIN_RECF_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 57 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 58 79 1 (POTENTIAL).
FT DOMAIN 80 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 131 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 132 150 3 (POTENTIAL).
FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 189 4 (POTENTIAL).
FT DOMAIN 190 219 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 220 242 5 (POTENTIAL).
FT DOMAIN 243 336 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 337 358 6 (POTENTIAL).
FT DOMAIN 359 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 7 (POTENTIAL).
FT DOMAIN 398 450 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 127 205 BY SIMILARITY.
FT LIPID 411 411 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 450 AA; 48755 MW; EF13BA8B5FAA857A CRC64;

Query Match 10.8%; Score 237.5; DB 1; Length 450;
Best Local Similarity 23.5%; Pred. No. 3e-09;
Matches 88; Conservative 71; Mismatches 149; Indels 67; Gaps 13;

NY6R_MOUSE STANDARD; PRT; 371 AA.

AC Q61212;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Neuropeptide Y receptor type 6 (NPY6-R) (Pancreatic polypeptide receptor 2) (PP2).

GN NPY6R OR PPR2 OR NPY5R.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129;

RX MEDLINE=96279200; PubMed=8663568;

RA Weinberg D.H., Sirinathsinghji D.J.S., Tan C.P., Shiao L.-L., Morin N., Rigby M.R., Heavens R.H., Rapoport D.R., Bayne M.L., Cascleri M.A., Strader C.D., Linemeyer D.L., Macneil D.J.; "Cloning and expression of a novel neuropeptide Y receptor."; J. Biol. Chem. 271:16435-16438(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/Sv;

RX MEDLINE=97066971; PubMed=8910373;

RA Gregor P., Feng Y., Decarr L.B., Cornfield L.J., McCaleb M.L.; "Molecular characterization of a second mouse pancreatic polypeptide receptor and its inactivated human homologue."; J. Biol. Chem. 271:2776-2778(1996).

CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE RANK ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS NPY > PYY > NPY (2-36) = [LEU-31, PRO-34] NPY > NPY (13-36) > PP. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: KIDNEY AND DISCRETE REGIONS OF THE HYPOTHALAMUS INCLUDING THE SUPRACHIASMATIC NUCLEUS, ANTERIOR HYPOTHALAMUS, BED NUCLEUS STRIA TERMINALIS, AND THE VENTROMEDIAL NUCLEUS.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.

CC -!- CAUTION: Was originally called NPY5-R.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U58367; AAB18624.1; -

DR EMBL: U59430; AAB19188.1; -

DR MGD; MGI:1098590; Npy6r.

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 35 57 1 (POTENTIAL).

FT DOMAIN 58 69 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 70 90 2 (POTENTIAL).

FT DOMAIN 91 110 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 111 132 3 (POTENTIAL).

FT DOMAIN 133 152 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 153 174 4 (POTENTIAL).

FT DOMAIN 175 208 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 209 230 5 (POTENTIAL).

FT DOMAIN 231 261 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 262 284 6 (POTENTIAL).

FT DOMAIN 285 297 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 298 321 7 (POTENTIAL).

FT DOMAIN 322 371 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DISULFID 109 196 BY SIMILARITY.

FT LIPID 336 PALMITATE (POTENTIAL).

SQ SEQUENCE 371 AA; 42713 MW; E4AABB987CEB74B7 CRC64;

Query Match 10.9%; Score 239; DB 1; Length 371;

Best Local Similarity 23.3%; Pred. No. 1.9e-09;

Matches 90; Conservative 60; Mismatches 109; Indels 128; Gaps 17;

QY 52 IVLGIGSLNLTVLVLYCMKNSNLTMTNHLVLDVILCVGCIPLTIVILLLS--LES 109

DB 46 LIMGIFGNLSLIIIFKKQREAGQNVNLTIANLSLSLILVCMCIPTVIYITLMDHWVF 105

QY 110 NTALICCFEACVSFASVSTAI-NVFAITLDYDISVKPANRILTMGRAYMLMISWIFS 168

DB 106 NT---MCKLTSYVQSVSVSIFSLVLIATERYQLIVNPRGKPRVAHAYWGIIWLIS 162

QY 169 -----FFSFLI---PF-----IEVNFPSLQSG--NTWENK--TLLCVSTN----EYTEL 207

DB 163 LTLISPLFLSYHLTNEPFFHNLSTLDIYTHQVACVEIWPCKLNQLLSTSLFMLQYFVPL 222

QY 208 GMYVHLLVQPIFFFTVVMVLTITKLOALNIRIGTFSTGQKKARKKKTISLTQHE 267

DB 223 GF-----LICYLKIVLCI-----RRRT----- 240

QY 268 ATDMSSSGGRNVFVGRTSVSVIIALRRVAKRHRERRER---QKRVFRMSLLIISTFLL 324

DB 241 -----RQVDRKKNKSLNENKRVNVMVLSIVVTFGA 272

QY 325 CWTPISVLNT-----TILCLGFSDDLVLKRLCLFLVMAYGTTTTPHLLIYFTQKPKVL 378

DB 273 CWLPLNIFNFIDWYHEMLMSCHDLV--FVVCGLI-AMVSTCINPLFYGLNKNFKQDL 329

QY 379 KSKMKRVVSVSEADPLPNNAVHNSW 405

DB 330 M-----MLIHWCW 337

RESULT 9

NY6R_RABIT STANDARD; PRT; 371 AA.

AC P79217;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Neuropeptide Y receptor type 6 (NPY6-R).

GN NPY6R.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Japanese white; TISSUE=Skeletal muscle;

RX MEDLINE=97066888; PubMed=8910290;

RA Matsumoto M., Nomura T., Momose K., Ikeda Y., Kondou Y., Akiho H., Togami J., Kimura Y., Okada M., Yamaguchi T.; "Inactivation of a novel neuropeptide Y/peptide YY receptor gene in primate species."; J. Biol. Chem. 271:27217-27220(1996).

CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.

CC -----

FT DISULFID 114 196 BY SIMILARITY.
FT LIPID 387 387 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 428 AA; 47641 MW; A6E8FABDA805E610 CRC64;
Query Match 11.5%; Score 253; DB 1; Length 428;
Best Local Similarity 22.8%; Pred. No. 2.5e-10;
Matches 95; Conservative 85; Mismatches 132; Indels 104; Gaps 18;
QY 15 SNITVRDIDDINTNMV-----QPLSYPLSFQVSLTGLMLEIVLGLSN 59
Db 12 SNITPPCELGLENETFLCDQPRPSKEWQPAVQLLYSLIFLLSV-----LGNT 60
QY 60 LTVLVLYCMKNSLINSVNIITMNLHVDLVICVGPITVITVILLLSLESNTALICCPHE 119
Db 61 LVITVL--IRNKRMTVNIILSLAVSDMLCLFCMPFNILPNLLKDFIEGSAVC--K 115
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-ITMGRAVLMISWIFSF----- 169
Db 116 TTTVPMGTSVSVTFNLVAISLERYGAICKPLQSRVMTKSHALKVIAATWCLSFITMP 175
QY 170 ---ESFLIPFTEVNFSSQSGNTWENKTL-LC--VSTNEYTELGMVYHLLVQPIFFFT 223
Db 176 YPIYSNLVFFTKNN-----NOTANMCRFLPNV---MOQSWHTFLLLFLIP 221
QY 224 VVMLITVTKLQALNIRIGTRFSTGQKKARKKTKTISLTT-OHEATD----- 270
Db 222 GIVMVAAGLI--SLELQGIKFASQKSAKERKPTSTSSGKYEDSGCYLQKTRPPRK 279
QY 271 -----MSOSSGGRNVFVGRSVSVIIALRAVKKHRRERQKRVFMSLLIISTFLLC 325
Db 280 LELQSLSTGSSSRANRIRSNSSAALMA-----KKRVIRMLIVIVVLPFLC 325
QY 326 WTPISLVNTTILCLGPSLLVLRCL-----FLVMAYGTFIHFPLLYAFTRQFQ 375
Db 326 WMPISANA---WRAYDTASAERLSGTPISIFLLSYTSVCVNPPIIYCFMKNKRF 378
RESULT 7
CCKR_XENLA STANDARD; PRT; 453 AA.
ID CCKR_XENLA
AC P0031;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cholecystokinin receptor (CCK-XLR).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96319796; PubMed=8700154;
RA Schmitz F., Pratt D.S., Wu M.-J., Kolakowski L.F. Jr., Beinhorn M.,
RA Kopin A.S.;
RT "Identification of cholecystokinin-B/gastrin receptor domains that
RT confer high gastrin affinity: utilization of a novel xenopus laevis
RT cholecystokinin receptor.";
RL Mol. Pharmacol. 50:436-441(1996).
CC -1- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. THIS RECEPTOR MEDIATES ITS
CC ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM. HAS HIGH
CC AFFINITY FOR CCK-8 AND LOW AFFINITIES FOR GASTRIN-17-I, CCK-4, AND
CC UNSULFATED CCK-8.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: BRAIN AND STOMACH.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HAS EQUAL SIMILARITY TO TYPE A AND B CHOLECYSTOKININ MAMMALIAN
CC RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; U49258; AAB09052.1; --
DR HSP; P02699; I788
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 64
FT TRANSMEM 65 94
FT TRANSMEM 95 104
FT TRANSMEM 105 131
FT TRANSMEM 132 142
FT TRANSMEM 143 164
FT TRANSMEM 165 184
FT TRANSMEM 185 205
FT TRANSMEM 206 237
FT TRANSMEM 238 261
FT TRANSMEM 262 343
FT TRANSMEM 344 364
FT TRANSMEM 365 379
FT TRANSMEM 380 403
FT TRANSMEM 404 423
FT DISULFID 141 223
FT LIPID 401 401
FT CARBOHYD 9 9
FT CARBOHYD 22 22
FT CARBOHYD 30 30
FT CARBOHYD 35 35
FT CARBOHYD 39 39
FT CARBOHYD 401 401
SQ SEQUENCE 453 AA; 51157 MW; 06217927B7482678 CRC64;
Query Match 11.3%; Score 248.5; DB 1; Length 453;
Best Local Similarity 22.7%; Pred. No. 5.4e-10;
Matches 93; Conservative 84; Mismatches 172; Indels 61; Gaps 14;
QY 32 QPLSYPLSFQVSLTGLMLEIVLGLSNLTIVLVLYCMKNSLINSVNIITMNLHVDVII 91
Db 60 KPSPDLNLWRIVMYSVIFLLSVFGNTLIIVLMNKR--LRITNSFLLSLALSILMV 117
QY 92 CVGCIPTIVILLLSLENTALICFCHEACVSFASVAINVF---AITLDYDISVKP- 147
Db 118 AVLCPFTLIPNLMEFIFGEVIC---RAAYFVGLSVSVTFNLVAISERYSAICNPL 174
QY 148 ANRI-LTMGRAVLMISWIFSFIFLPIFTEVN---FFSL-----QSGNTWENKTL 196
Db 175 KSRVQWTRSHAYRIAATAWTLSSIT-IMPILVYNKTVTFPMKRRVGHQCRVWPSK--- 230
QY 197 CVSTNEYTELGMVYHLLVQPIFFFTVVMVLYTKLQALNIRIGTRFSTGQKKARK 256
Db 231 -----QQQAWYVLLLTIFLPGVVMIVAYGLISRELYGIQFEMDLNKAHKK 281
QY 257 KKTISLTQHEATD-----MSOSSGGRNVFVGRSVSVIIALRRA-VKKHRRERQKRV 311
Db 282 NGVSTPTIPSGDEGDCYIQVTKRRNTMEMSTLTPSVCTKMDRARNINSEAKLMKRV 341
QY 312 FMSLLIISTFLLCWTPISLVNT--TILCLGPSDLLVLRCLCF-LVMAYGTFIHFPLLYA 368
Db 342 IRMLIVIVAMFFICWMPIFVANTWKADELAFNTLTGAPISFIIHLLSYTSACVNPLIYC 401
QY 369 FTROKFOKLKSKKKRVSVIVEADPLPNNVAINHNSWIDPKRNKKTIPED 418
Db 402 FNNRRFRKALGTF-----SSCIKPCRNFRDTE 431
RESULT 8

```
FT TRANSMEM 366 399 7 (POTENTIAL).
FT DOMAIN 330 444 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 129 211 BY SIMILARITY.
FT LIPID 403 403 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 444 AA; 49657 MW; B435BE75052FB11 CRC64;

Query Match
Best Local Similarity 23.8%; Pred. No. 3.2e-11;
Matches 102; Conservative 85; Mismatches 131; Indels 111; Gaps 20;

QY 6 ILEINNOSESNITVRDIDDINTNMV-----QP-LSYPLSFQVSLTGFLMLEIVLGLGNSN 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 VVDSLLMNGSNITPCPELGTNETLFCLDQPOPSKEWQSAQILLYSIIFLLSVLGLNTLV 77
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 LVVLVLYCMKNSLINSVSNITMNLHVLVDIICVGCIPUTIVILLLSLESNTALICCFHE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 ITVLI-----RNKRMRVTWTFIFLLSLAVSDMLCLFCMPENLIPNLLKDFIFGSAVC---K 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 ACVSFASVSTAINVF---AITLDRDISVKP-ANRI-LTMGRAVMLMISIWIFSF----- 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 TTYTFMGTSVSTFNLVAISLIERGAICRPLQSRVQWQKSHALKVIAATWGLSFTIMTP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 ---FSLPIPFIEIVN-----FFSLQSG---NTWENKTLVYSTNBYTELGMYYHLVLVQ 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 YPIYSNLVPFTKNNQNTANMCRFLPDSAMQSQWOTFLLI----- 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 217 IFIEFTVVMVLTITKILQALNIRIGTFSTGCKKARKKK-TISLTQHEATD----- 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 --LFLPLGIVMVAVGLI--SLEVOGIRKFDASQKAKEKPGSTGSTRYEDSDGCIYQ 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 -----MSQSSGGRNVVFGVTSVSV--IIALRRVAKRRREROKRVPRMSL 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 KSRPPKLEQLQSSGSGSR-LNRRSSSSAANLIA-----KKRVIRMLI 332
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 LIISTFLLCWTPISVLNT-----TI-----LCIGPSDLLVKRLCFLVWAYGTTFPHPL 366
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 333 VIVVLFCLWMPIFSANAWRAYDTVSAERHLSGTPISFI-----LLLSYTSNCVNPII 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 YAFTRQKQF 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 386 YCFMNRFR 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
CCKR_HUMAN
ID CCKR_HUMAN STANDARD; PRT; 428 AA.
AC P32338;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN CCKAR OR CCKRA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Gall bladder;
RX MEDLINE=93277552; PubMed=8503909;
RA Ulrich C.D., Ferber I., Hollicky E., Hadac E., Buell G.,
RA Miller L.J.;
RT "Molecular cloning and functional expression of the human gallbladder
cholecystokinin A receptor.";
RL Biochem. Biophys. Res. Commun. 193:204-211(1993).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=93343941; PubMed=8343165;
RA Wank S.A., de Weerth A., Pilegna J.R., Huppi K.;
RT "Molecular cloning, functional expression and chromosomal
```

```
RT localization of the human cholecystokinin type A receptor.";
RL Biochem. Biophys. Res. Commun. 194:811-818(1993).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96029343; PubMed=7557108;
RA Miller L.J., Hollicky E.L., Ulrich C.D., Wieben E.D.;
RT "Abnormal processing of the human cholecystokinin receptor gene in
association with gallstones and obesity.";
RL Gastroenterology 109:1375-1380(1995).
[4]
SEQUENCE FROM N.A.
RP TISSUE=Peripheral blood leukocytes;
RX MEDLINE=20145045; PubMed=10682840;
RA Funakoshi A., Miyasaka K., Matsumoto H., Yamamori S., Takiguchi S.,
RA Kataoka K., Takata Y., Matsuse K., Kono A., Shimokata H.;
RT "Gene structure of human cholecystokinin (CCK) type-A receptor: body
fat content is related to CCK type-A receptor gene promoter
polymorphism.";
RL FEBS Lett. 466:264-266(2000).
CC -!- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. HAS A 1000 FOLD AFFINITY
FOR CCK RATHER THAN FOR GASTRIN. IT MODULATES FEEDING AND
DOPAMINE-INDUCED BEHAVIOR IN THE CENTRAL AND PERIPHERAL NERVOUS
SYSTEM. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G
PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; L13605; AAA35659.1; -
DR EMBL; L19315; AAA02819.1; -
DR EMBL; U23430; AAA91123.1; -
DR EMBL; U23427; AAA91123.1; JOINED.
DR EMBL; U23428; AAA91123.1; JOINED.
DR EMBL; U23429; AAA91123.1; JOINED.
DR EMBL; D85606; BAA90879.1; -
DR PIR; JN0590; JN0590.
DR PIR; JN0692; JN0692.
DR HSP; P02699; 1F88.
DR Genew; HGNC:1570; CCKAR.
DR MIM; 118444; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 42 67 1 (POTENTIAL).
FT DOMAIN 68 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 104 2 (POTENTIAL).
FT DOMAIN 105 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 137 3 (POTENTIAL).
FT DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 234 5 (POTENTIAL).
FT DOMAIN 235 313 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 314 334 6 (POTENTIAL).
FT DOMAIN 335 349 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 350 373 7 (POTENTIAL).
FT DOMAIN 374 428 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; AF015963; AAC07949.1;
CC EMBL; AF015959; AAC07949.1; JOINED.
CC EMBL; AF015960; AAC07949.1; JOINED.
CC EMBL; AF015961; AAC07949.1; JOINED.
CC EMBL; AF015962; AAC07949.1; JOINED.
CC EMBL; D85605; BAA20068.1;
CC HSSP; P02699; 1F88.
CC MGD; MGI:99478; Cckar.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECF_F1_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECF_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 42 67 1 (POTENTIAL).
FT DOMAIN 68 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 104 2 (POTENTIAL).
FT DOMAIN 105 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 137 3 (POTENTIAL).
FT DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 234 5 (POTENTIAL).
FT DOMAIN 235 321 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 322 342 6 (POTENTIAL).
FT DOMAIN 343 357 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 358 381 7 (POTENTIAL).
FT DOMAIN 382 436 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 114 196 BY SIMILARITY.
FT LIPID 395 395 PALMITATE (BY SIMILARITY).
FT DOMAIN 258 264 POLY-GLY.
FT DOMAIN 265 268 POLY-SER.
CC SEQUENCE 436 AA; 48436 MW; EBB5D8453B792F79 CRC64;

Query Match 12.2%; Score 267; DB 1; Length 436;
Best Local Similarity 23.8%; Pred. No. 3e-11;
Matches 104; Conservative 85; Mismatches 128; Indels 120; Gaps 20;

QY 6 ILEINMQSESNTVRDDIDDTNNMY-----OP-LSYPLSFQVSLTGFLMLEIVLGLSN 59
DB 3 VVDSLLMGNSNITPPCELGLENETLFCLDQPPQSKEMQSAVOILLYSFILLSVLGNTLV 62
QY 60 LTVLLVLCMKNSLINSNIITMNLRLVDVIVCGICPLTIVILLLESNTALICCPHE 119
DB 63 ITVLI-----RNKMTVTNIFLLSLAVSDMLCLFCMFNPLNLLKDFIGSAVC---K 115
QY 120 ACVSPASVSTAINVF---AITLDYDISVKP-ANKI-ITMGRAVLMISIMWFSF----- 169
DB 116 TTTVMGTSVSVSTNLVAISLERYGAICRPLQSRVQWTKSHALKVIAATWCLSTIWP 175
QY 170 ---FSELPFFTEVN-----FFSLQSG---NTWENKLLCVSTNEYTYELGMYYHLVQ 216
DB 176 YPIYSLNLPVFTKNNQNTAMCRFLPDSAMQSQWTFLLI-----SLTTOHEA 268
QY 217 IPIPFETVVVMLITVTKILQALNIRIGRFSTGQKKKARKKTI-----SLTTOHEA 268
DB 217 --LFLIPGVVWVAIGLI--SLELTQGIKFDASQKSKAKRLSSGGGGSSSRVD 272
QY 269 TD-----MSQSSGGRRNVVGVRTSVS--IIALRAVRRHRRRRRQ 308
DB 273 SDGCYLOKSRPRKLEAQLSTSSSGGR--INRISSGAANLIA-----K 316

```

```

QY 309 KRVRMSLLIISTFLCWTPIISVLNT-----TI-----LCLGPSDLLVYKLRCLFLVMAYG 358
DB 317 KRVRMLIVIVVFLCWNPIFSANAWRAYDVTSAEKHLSGTPIISFI-----LLLSYT 369
QY 359 TTIFPLLYAFTRQKFPQ 375
DB 370 SSCVNPPIIYCFMKNRFR 386

RESULT 5
CCRR_RAT STANDARD; PRT; 444 AA.
AC P30551;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN CCKAR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 111-158; 270-314 AND 392-402.
RC TISSUE-Pancreas;
RX MEDLINE-92212981; PubMed-1313582;
RA Wank S.A., Harkins R., Jensen R.T., Shapira H., de Weerth A.,
RA Slattery T.;
RT "Purification, molecular cloning, and functional expression of the
RT cholecystokinin receptor from rat pancreas.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3125-3129(1992).
CC -1- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. HAS A 1000 FOLD AFFINITY
CC FOR CCK RATHER THAN FOR GASTRIN. IT MODULATES FEEDING AND
CC DOPAMINE-INDUCED BEHAVIOR IN THE CENTRAL AND PERIPHERAL NERVOUS
CC SYSTEM. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G
CC PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PANCREAS AND BRAIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

QY 327 TPISLVNT-----TI-----LCIGPSDLLVLRCLCELMVAYGTTIFPHLLYATFTRQKFQ 375
 DB 329 MPISANAWRAYDTVAERHLSGTPISFI-----LLSYTSSCVNPIIYCFMNRFR 380

RESULT 3
 CCKR_RABIT STANDARD; PRT; 427 AA.

AC 097772;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
 GN CCKAR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=95002144; PubMed=7918628;
 RA Reuben M., Rising L., Prinz C., Hersey S., Sachs G.;
 RT "Cloning and expression of the rabbit gastric CCK-A receptor.";
 RL Biochim. Biophys. Acta 1219:321-327(1994).
 CC -!- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. HAS A 1000 FOLD AFFINITY
 CC FOR CCK RATHER THAN FOR GASTRIN. IT MODULATES FEEDING AND
 CC DOPAMINE-INDUCED BEHAVIOR IN THE CENTRAL AND PERIPHERAL NERVOUS
 CC SYSTEM. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G
 CC PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U86601; AAD11547.1; -
 DR HSP; P02699; I188.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Lipoprotein; Palmitate.
 FT DOMAIN 1 41
 FT TRANSMEM 42 67
 FT DOMAIN 68 77
 FT TRANSMEM 78 104
 FT DOMAIN 105 115
 FT TRANSMEM 116 137
 FT DOMAIN 138 157
 FT TRANSMEM 158 178
 FT DOMAIN 179 210
 FT TRANSMEM 211 234
 FT DOMAIN 235 312
 FT TRANSMEM 313 333
 FT DOMAIN 334 348
 FT TRANSMEM 349 372
 FT DOMAIN 373 427
 FT CARBOHYD 10 10
 FT CARBOHYD 24 24
 FT CARBOHYD 190 190
 FT DISULFID 114 196
 FT LIPID 366 386
 SEQUENCE 427 AA; 47386 MW; 089FD10E2B86DB25 CRC64;

Query Match 12.4%; Score 272; DB 1; Length 427;
 Best Local Similarity 24.8%; Pred. No. 1.3e-11;
 Matches 94; Conservative 75; Mismatches 130; Indels 80; Gaps 16;

QY 32 QPLSYPLSFQVSLTGFLMLEIVLGLSGLNLTVLVLYCMKSNLINSVSIITMNLHVLVDVII 91
 DB 44 QILLYSLIFLLSV-----IGNTLIVITVL--IRNKRMTVTNIFLLSLAISDLML 90
 QY 92 CVCICPLTIV-IILLSESNALICCFHEACVSFASVS-TAINVFAITLDRDISVKP-A 148
 DB 91 CLFCMEFNLIPNLLKDFIFGSAL--CKTTYLMGTSVSYSTLNLVASIERGAICKPLQ 148
 QY 149 NRI-LFMGRVLMIMISINIFSP-----FSLPIPIEVNFFSLOSQGNWENKTLQVS 199
 DB 149 SRVQTKSHALKVIAATWCLSFALMTPYPIYSNLVPFTKTN-----NOTANMCRFLP 201
 QY 200 TNEYYTELGMWYHLLVQIPIFFFTVVMLTYTKILOALNIRIGTRFSGQKKARKKT 259
 DB 202 SD-----VMOQAWHTFLLLILFLIPGIVMVAYGMI--SLELYQGKIFDASQKSAKERKA 255
 QY 260 IS-----LTTQHEATDMSQSSGGRNVVFGVRTSVSVIIALRRVAKRHR 302
 DB 256 STGSGRFEDNDCYLORSKPTQLQLQQLSGGGGR--VSRIRSSSAATLMAK----- 307
 QY 303 ERRERQKRVFRSLIISTFLCWTPTISVLNTILICLPSDLLVLRCLC-----FLVMA 356
 DB 308 -----KRVIRMLWVIVLFFLCWMPIFSANA---WRAYDTVSAERRLSGTPISFILLLS 358
 QY 357 YGTTTFPHLLYATFTRQKFQ 375
 DB 359 YTSSCVNPIIYCFMNRFR 377

RESULT 4
 CCKR_MOUSE STANDARD; PRT; 436 AA.

AC 008786;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
 GN CCKAR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ; TISSUE=Liver;
 RX MEDLINE=97254481; PubMed=9099891;
 RA Takata Y., Takiguchi S., Takaoka K., Funakoshi A., Miyasaka K.,
 Kono A.;
 RA "Mouse cholecystokinin type-A receptor gene and its structural
 analysis.";
 RL Gene 187:267-271(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=97396148; PubMed=9245702;
 RA Lacourse K.A., Lay J.M., Swanberg L.J., Jenkins C., Samuelson L.C.;
 RT "Molecular structure of the mouse CCK-A receptor gene.";
 RL Biochem. Biophys. Res. Commun. 236:630-635(1997).
 CC -!- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. HAS A 1000 FOLD AFFINITY
 CC FOR CCK RATHER THAN FOR GASTRIN. IT MODULATES FEEDING AND
 CC DOPAMINE-INDUCED BEHAVIOR IN THE CENTRAL AND PERIPHERAL NERVOUS
 CC SYSTEM. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G
 CC PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

FT DOMAIN 178 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 5 (POTENTIAL).
FT DOMAIN 230 315 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 316 336 6 (POTENTIAL).
FT DOMAIN 337 370 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOMAIN 371 433 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 425 425 R -> C (IN REF. 2).
SQ SEQUENCE 433 AA; 49318 MW; 385868AC567C4DC1 CRC64;

Query Match 100.0%; Score 2192; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 5.2e-141;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCFSPLEINMQSESNITVRDIDDINTNMYQPLSYPLSFQVSLTGLFMLEIVLGLGSLN 60
Db 1 MCFSPLEINMQSESNITVRDIDDINTNMYQPLSYPLSFQVSLTGLFMLEIVLGLGSLN 60

Qy 61 TVLVLYCHKSNLINSVNIITMNLHLVDVLCVGCIPITIVILLLESNTALICCFHEA 120
Db 61 TVLVLYCHKSNLINSVNIITMNLHLVDVLCVGCIPITIVILLLESNTALICCFHEA 120

Qy 121 CVSPASVSTAINVPAITLDYDISVKPANRLTMTGRAVLMISITWIFSFELIPFIEVN 180
Db 121 CVSPASVSTAINVPAITLDYDISVKPANRLTMTGRAVLMISITWIFSFELIPFIEVN 180

Qy 181 FFSQSGNTWENKTLVYSTNEYTELGMVYHLLVQIPFFFTVVVMLITTKILQALNI 240
Db 181 FFSQSGNTWENKTLVYSTNEYTELGMVYHLLVQIPFFFTVVVMLITTKILQALNI 240

Qy 241 RIGTFSTGOKKARKKKTISLTTOHEATMSQSGGRNVFVGVTSVSVIALRAVYKR 300
Db 241 RIGTFSTGOKKARKKKTISLTTOHEATMSQSGGRNVFVGVTSVSVIALRAVYKR 300

Qy 301 HREREROKRVFRMSLLIISITFLCWTPISVLNTTILGSPDLLVKRLCFLVMAYGTT 360
Db 301 HREREROKRVFRMSLLIISITFLCWTPISVLNTTILGSPDLLVKRLCFLVMAYGTT 360

Qy 361 IFHPLLYAFTQKFKVLSKMKRRVSVIADPLPNNAVHNSWIDPKRKKITFEDE 420
Db 361 IFHPLLYAFTQKFKVLSKMKRRVSVIADPLPNNAVHNSWIDPKRKKITFEDE 420

Qy 421 IREKRLVPQVYTD 433
Db 421 IREKRLVPQVYTD 433

```

RESULT 2

```

CCKR_CAVPO STANDARD; PRT; 430 AA.
AC Q63931;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN CCKAR.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Pancreas, and Gall bladder;
RX MEDLINE=94106629; PubMed=7916580;
RA "Guinea pig gallbladder and pancreas possess identical CCK-A receptor
RT subtypes: receptor cloning and expression.";
RL Am. J. Physiol. 265:G1116-G1121(1993).
CC -1- FUNCTION: RECEPTOR FOR CHOLECYSTOKININ. HAS A 1000 FOLD AFFINITY
CC FOR CCK RATHER THAN FOR GASTRIN. IT MODULATES FEEDING AND
CC DOPAMINE-INDUCED BEHAVIOR IN THE CENTRAL AND PERIPHERAL NERVOUS

```

```

CC SYSTEM. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G
CC PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S68242; AAB29504.1; -
CC HSSP; P02699; 1F88.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODOPS
CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Lipoprotein; Palmitate.
CC DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 42 67 1 (POTENTIAL).
CC DOMAIN 68 77 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 78 104 2 (POTENTIAL).
CC DOMAIN 105 115 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 116 137 3 (POTENTIAL).
CC DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 158 178 4 (POTENTIAL).
CC DOMAIN 179 210 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 211 234 5 (POTENTIAL).
CC DOMAIN 235 315 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 316 336 6 (POTENTIAL).
CC DOMAIN 337 351 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 352 375 7 (POTENTIAL).
CC DOMAIN 376 430 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DISULFID 114 196 BY SIMILARITY.
CC LIPID 389 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 430 AA; 48210 MW; FC9F5D34032076C9 CRC64;

Query Match 12.9%; Score 282; DB 1; Length 430;
Best Local Similarity 24.1%; Pred. No. 2.8e-12;
Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;

Qy 15 SNITVRDIDDINTNMY-----QPLSYPLSFQVSLTGLFMLEIVLGLGSLN 59
Db 12 SNITVACELGFENETLFCLEDRPRPSKEWQPAVQILLYSLIFLLSV-----LGNT 60

Qy 60 LTVLVLYCHKSNLINSVNIITMNLHLVDVLCVGCIPITIVILLLESNTALICCFHE 119
Db 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCFCMFPNLPISLLKDFIGSAVC---K 115

Qy 120 ACVSFASVSTAINVFP---AITLDYDISVKP-ANRI-LTMGRAVLMISITWIFSF----- 169
Db 116 TTTYPMGTGSVSVSTENLVLSIERYGAIKCPQSRVWQTKSHALAKAVIAATWCLSTWTP 175

Qy 170 ---FSFLIPFTEVNFSLQSGNTWENKTLIC--VSTNEYTELGMVYHLLVQIPFFFTV 224
Db 176 YPIYSNLVPFTKNN---NOTGN-----MCRFLLPNDVMQOT---WHTFLLLILFLIPG 222

Qy 225 VVMLITYTKILQALNIRIGTRFSTGOKKARKKKTIS-----LTTOH----- 266
Db 223 IVMMVAYGLI--SLEYQGIKFDAIQKSAKERKTSTGSGPMEDSDGCVLQKSRHPRKL 280

Qy 267 EATDMSQSGGRNVFVGVTSVSVIALRAVYKRERREROKRVFRMSLLIISITFLC 326
Db 281 ELRQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVLVFLCW 328

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:59:28 ; Search time 14 Seconds
(without alignments)
1282.804 Million cell updates/sec

Title: US-09-845-721-2

Perfect score: 2192

Sequence: 1 MCFSPILEINQMSNSNIVR.....ITFDESEIREKRLVPQVYTD 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 2192 | 100.0 | 433 | 1 GP22_HUMAN | Q99680 homo sapien |
| 2 | 282 | 12.9 | 430 | 1 CCRX_CAVPO | Q63931 cavia porce |
| 3 | 272 | 12.4 | 427 | 1 CCRX_RABIT | Q97772 oryctolagus |
| 4 | 267 | 12.2 | 436 | 1 CCRX_MOUSE | O08786 mus musculus |
| 5 | 266.5 | 12.2 | 444 | 1 CCRX_RAT | P30551 rattus norv |
| 6 | 253 | 11.5 | 428 | 1 CCRX_HUMAN | P32238 homo sapien |
| 7 | 248.5 | 11.3 | 453 | 1 CCRX_XENLA | P70031 xenopus lae |
| 8 | 239 | 10.9 | 371 | 1 NY6R_MOUSE | Q61212 mus musculus |
| 9 | 237.5 | 10.8 | 371 | 1 NY6R_RABIT | P79217 oryctolagus |
| 10 | 237.5 | 10.8 | 450 | 1 GASR_PRANA | P30796 praomys nat |
| 11 | 226 | 10.3 | 453 | 1 GASR_MOUSE | P56481 mus musculus |
| 12 | 222.5 | 10.2 | 452 | 1 GASR_RAT | P30553 rattus norv |
| 13 | 222 | 10.1 | 422 | 1 SH1A_HUMAN | P08908 homo sapien |
| 14 | 218.5 | 10.0 | 370 | 1 SH5B_MOUSE | P31387 mus musculus |
| 15 | 218 | 9.9 | 447 | 1 GASR_HUMAN | P32239 homo sapien |
| 16 | 214.5 | 9.8 | 454 | 1 GASR_BOVIN | P79266 bos taurus |
| 17 | 213.5 | 9.7 | 357 | 1 SH5A_RAT | P35364 rattus norv |
| 18 | 213.5 | 9.7 | 452 | 1 GASR_RABIT | P46627 oryctolagus |
| 19 | 213.5 | 9.7 | 509 | 1 5HT_LYMST | Q25414 lymnaea sta |
| 20 | 213 | 9.7 | 453 | 1 GASR_CANFA | P30552 canis famli |
| 21 | 212.5 | 9.7 | 460 | 1 OX2R_RAT | P56719 rattus norv |
| 22 | 211.5 | 9.6 | 478 | 1 OX2R_HUMAN | Q9umh6 homo sapien |
| 23 | 211 | 9.6 | 521 | 1 OPN4_MOUSE | Q9qxx9 mus musculus |
| 24 | 210.5 | 9.6 | 373 | 1 GP45_MOUSE | Q9eqq4 mus musculus |
| 25 | 210.5 | 9.6 | 515 | 1 A1AB_RAT | P15823 rattus norv |
| 26 | 209.5 | 9.6 | 357 | 1 SH5A_HUMAN | Q47898 homo sapien |
| 27 | 209.5 | 9.6 | 455 | 1 YXAS_CAEEL | Q18179 caenorhabdi |
| 28 | 208.5 | 9.5 | 370 | 1 SH5B_RAT | P35365 rattus norv |
| 29 | 208 | 9.5 | 383 | 1 NY1R_CAVPO | Q9wvd0 cavia porce |
| 30 | 208 | 9.5 | 383 | 1 NY1R_PIG | O02835 sus scrofa |
| 31 | 208 | 9.5 | 515 | 1 A1AB_MESAU | P18841 mesocricetu |
| 32 | 207 | 9.4 | 519 | 1 A1AB_HUMAN | P35368 homo sapien |
| 33 | 206 | 9.4 | 421 | 1 SH1A_MOUSE | Q64264 mus musculus |

| | | | | | |
|----|-------|-----|-----|--------------|---------------------|
| 34 | 205 | 9.4 | 423 | 1 SH1A_FUGRU | O42385 fugu rubrip |
| 35 | 204 | 9.3 | 384 | 1 NY1R_HUMAN | P25929 homo sapien |
| 36 | 204 | 9.3 | 457 | 1 DBDR_XENLA | P42290 xenopus lae |
| 37 | 203.5 | 9.3 | 326 | 1 AA1R_BOVIN | P28190 bos taurus |
| 38 | 203.5 | 9.3 | 446 | 1 DADR_DIDMA | P42288 didelphis m |
| 39 | 203 | 9.3 | 382 | 1 NY1R_CANFA | O02813 canis famli |
| 40 | 203 | 9.3 | 382 | 1 NY1R_MOUSE | Q04573 mus musculus |
| 41 | 203 | 9.3 | 382 | 1 NY1R_RAT | P21555 rattus norv |
| 42 | 202.5 | 9.2 | 444 | 1 OX2R_CANFA | Q9tup7 canis famli |
| 43 | 202 | 9.2 | 326 | 1 AA1R_CANFA | P11616 canis famli |
| 44 | 202 | 9.2 | 451 | 1 DADR_XENLA | P42289 xenopus lae |
| 45 | 201.5 | 9.2 | 357 | 1 SH5A_MOUSE | P30966 mus musculus |

ALIGNMENTS

RESULT 1

| ID | GP22_HUMAN | STANDARD; | PRT; | 433 AA. |
|----|--|-----------|------|---------|
| AC | Q99680; O14554; | | | |
| DT | 01-NOV-1997 (Rel. 35, Created) | | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | Probable G protein-coupled receptor GPR22. | | | |
| GN | GPR22. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=97225799; PubMed=9073069; | | | |
| RA | O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R., | | | |
| RA | Heng H.H.Q., Kolakowski L.F. Jr., Lynch K.R., George S.R.; | | | |
| RT | "Cloning and chromosomal mapping of four putative novel human | | | |
| RL | Gene 187:75-81(1997). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Gattung S.; | | | |
| RL | Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases. | | | |
| CC | -!- FUNCTION: ORPHAN RECEPTOR. | | | |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein. | | | |
| CC | -!- TISSUE SPECIFICITY: IN THE BRAIN REGIONS FRONTAL CORTEX, CAUDATE, | | | |
| CC | PUTAMEN, AND THALAMUS; NOT IN PONS, HYPOTHALAMUS, AND HIPPOCAMPUS. | | | |
| CC | -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; U66581; AAC51304.1; -- | | | |
| DR | EMBL; AC002381; AAB63815.1; -- | | | |
| DR | Genew; HGNC:4477; GPR22. | | | |
| DR | MM; 601910; | | | |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. | | | |
| DR | Pfam; PF00001; 7tm1; 1. | | | |
| DR | PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG. | | | |
| DR | PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1. | | | |
| KW | G-protein coupled receptor; Transmembrane; Glycoprotein. | | | |
| FT | DOMAIN 1 45 | | | |
| FT | TRANSMEM 46 66 | | | |
| FT | DOMAIN 67 85 | | | |
| FT | CYTOPLASMIC (POTENTIAL). | | | |
| FT | TRANSMEM 86 106 | | | |
| FT | DOMAIN 107 115 | | | |
| FT | EXTRACELLULAR (POTENTIAL). | | | |
| FT | TRANSMEM 116 136 | | | |
| FT | DOMAIN 137 156 | | | |
| FT | CYTOPLASMIC (POTENTIAL). | | | |
| FT | TRANSMEM 157 177 | | | |
| FT | 4 (POTENTIAL). | | | |

A:Note: authors translated the codon MET for residue 427 as Ile, and ACC for residue 440
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

```
Query Match          9.9%; Score 216; DB 2; Length 450;
Best Local Similarity 22.0%; Pred. No. 7.3e-10;
Matches 99; Conservative 77; Mismatches 181; Indels 92; Gaps 18;

QY 8 EINMQSESNIVRRDIDDINTNMQOPLSYPLSFQVSLTG-FLMLEIVLGLGSNLTVLVLY 66
Db 5 DTTMDGEGLLVERDS-----SFRI-LTGCFLLSLILSTLLGNTLVCAAV 47
QY 67 CMKSNLINSVSNIIIMNLHVLDDVIVCGICPLTIVILLLESNTALICCPHEAC---VS 123
Db 48 IPRHLRSKVNFVISLAVSDLLVAVLVMFKAV-----SEIAGFWPGSFCNIWVA 100
QY 124 F---ASVTAINVFAITLDRYDISVKP--ANRILTMGRAVLMISI-WIFSFFSELIPIFI 177
Db 101 FDMCSTASILNLCVISVDRYWAISSPFYERKMT-PKAAFTIMISVANTLSVLISFIP-V 158
QY 178 EVNFFSLOSGNTWE-NKTLICVSTNEYYTELGMYYHLLVQIPIFFFTVVVMLITYTKILQ 236
Db 159 QLNWHKATTTSFLDLNASLOGISMDNCDSLNRMYAISSSLISFYIPVAIMIVTYTRIYR 218
QY 237 ALNIRIGTRFSTGQKKKARKKKTISLTQHEATDMSQSSGGGRNVVGVRTSVSVIIALRR 296
Db 219 IAKOI-RRISALERAHVAKNCQNTSGNRRSMDCCQPESNFKMSF----- 263
QY 297 AVKRRERERQKRVFRMSLLIISTFLLCWTPISVLNTTI-----LCLGPS 342
Db 264 -----KRETKVLKTLVIMGVFVCCWLPFFVLNCMIPPCEPTQPSKGAEEFCINST 314
QY 343 DLLVKLRCLFLVMAYGTTIFHPLLYAFT---RQKQKVLKS-KMKKRVVSIVEADPLPNN 398
Db 315 TFDV-----FIWFGWANSLNPIIYAFNADFKAFTLLGCRYLCPSMGNAIETVSINN 369
QY 399 AVIHNSWIDPKRNKKTTFEDSEIREKRLV 427
Db 370 GAVFSSQHEPK-----GSSPESNLV 390
```

Search completed: December 4, 2002, 16:02:47
Job time : 24 secs

A:Accession: JCI1352
A:Molecule type: mRNA
A:Residues: 1-447 <PIS>
A:Cross-references: GB:L04473; NID:g179997; PIDN:AAA35660.1; PID:g179998
A:Experimental source: brain, gastric
R:Lee, Y.M.; Beinborn, M.; McBride, E.W.; Lu, M.; Koliakowski Jr., L.F.; Kopin, A.S.
J. Biol. Chem. 268, 8164-8169, 1993
A:Title: The human brain cholecystokinin-B/gastrin receptor. Cloning and characterization
A:Reference number: A46645; MUID:93216795; PMID:7681836
A:Accession: A46645
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-447 <LEE>
A:Cross-references: GB:L08112; NID:g306488; PIDN:AAA35657.1; PID:g306489
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:129156)
R:Song, I.; Brown, D.R.; Wiltshire, R.N.; Gantz, I.; Trent, J.M.; Yamada, T.
Proc. Natl. Acad. Sci. U.S.A. 90, 9085-9089, 1993
A:Title: The human gastrin/cholecystokinin type B receptor gene: alternative splice dono
A:Reference number: A48262; MUID:94022320; PMID:8415658
A:Accession: A48262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-447 <RES>
A:Cross-references: GB:L10822; NID:g406075; PIDN:AAC37528.1; PID:g406076
C:Genetics:
A:Gene: GDB:CKBR
A:Cross-references: GDB:136457; OMIM:118445
A:Map position: 11p15.5-11p15.4
A:Introns: 51/1; 135/1; 218/2; 271/1
C:Superfamily: neurokinin 1 receptor
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
F:57-61/Domain: transmembrane #status predicted <TM1>
F:91-116/Domain: transmembrane #status predicted <TM2>
F:131-150/Domain: transmembrane #status predicted <TM3>
F:171-192/Domain: transmembrane #status predicted <TM4>
F:219-243/Domain: transmembrane #status predicted <TM5>
F:334-354/Domain: transmembrane #status predicted <TM6>
F:369-392/Domain: transmembrane #status predicted <TM7>
F:7,30,36/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:82,154,300,442/Binding site: phosphate (Ser) (covalent) #status predicted
F:127-205/Disulfide bonds: #status predicted
F:321/Binding site: phosphate (Thr) (covalent) #status predicted
Query Match 9.9%; Score 218; DB 2; Length 447;
Best Local Similarity 22.8%; Pred. No. 5e-10;
Matches 88; Conservative 70; Mismatches 136; Indels 92; Gaps 15;
QY 36 YPLSFQVSLTGFLMLELVLGLGSLNVLVLYCMKSNLINSVNIITMNLHVLVDVITCVGC 95
DB 61 YAVIFLMSGNNLLIIVLGL-----SRLRTVTNAFLLSLAVSDLLAVAC 107
QY 96 IPTIVILLLESNTALICCFHEACVSF-----ASVSTAINVFATLDRYDISVXP--A 148
DB 108 MPFTLPNLMGTIFGVVIC-----KAVSYLMGVSVSVST--LSLVATALERYSAICRPLQA 162
QY 149 NRILTMGRAYMLMISWIFSFSLPIEIVNF-----LQSGNTWENKTLICVSTNE 202
DB 163 RVNQTRSHARVIVATWLLSGL-LMPYPYVTVVQVGPRLVCVHRWPS----- 211
QY 203 YYTELGMYYHLLVQIPIFFTVVVMILTYTKILOALNIRIGTRF-----STGQKKARKK 257
DB 212 --ARVQTWSVLLLLLLFFIPGVVMAYVGLISREL--YLGRLFDGSDSDSQSRVRNQ 267
QY 258 KTIISLTQH-----EATDNSQSGGRNVF-----GVRTSVSVIIA 293
DB 268 GLPQAVHQNGCRPEPTGAVGEDSGCVQLPRSRPALELTALPAGGSGSRPTQAKLLA 327
QY 294 LRRAVKRRRERQRKRVFRMSLLIISTFLCWTPIPSVLNLTICLGP--DLLVKRLC 351
DB 328 -----KKRVEMLLIVVFLCWLPPYSANTWRAFDGPGGAHRLSGAPIS 373
QY 352 FL-VMAYGTTIFHPLLYAFTRQFKQ 376

DB 374 FIHLLSYASACVNPPLYVCFMHRFRQ 399
RESULT 14
H87963
protein Y54E2A.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: H87963
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: H87963
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <STO>
A:Cross-references: GB:chr_I; PIDN:CAA21687.1; PID:g3881109; GSPDB:GN00019; CESP:Y54E
C:Genetics:
A:Gene: Y54E2A.1
A:Map position: 1
Query Match 9.9%; Score 217; DB 2; Length 580;
Best Local Similarity 22.2%; Pred. No. 7.9e-10;
Matches 85; Conservative 79; Mismatches 175; Indels 44; Gaps 15;
QY 40 FOVSLTGFLMLEIVL--GLGSNLTVLVLYCMKSNLINSVNIITMNLHVLVDVITCVGCIP 97
DB 7 FEIALPGFLYTLVFLVGTIGNSMVIFVNNFRK--MRNVTNIFLASLSTADLCILWFQVP 64
QY 98 LTVIVILLLESNTALICCFHEACVS-----PASVSTAINVFATLDRYDISVVKPANRI 151
DB 65 ----IMFMKYNMTWSMGRF--ACYSVHYIQQFTCPGCVLTMTMISFERELATAYPMRNI 118
QY 152 --LTMGRAYMLMISWIFSFSLPIEIVNF---FSLQSGNT--WENKTLICVSTNEYTT 205
DB 119 WFSIGRAKKVILLIWMSSAVLAPTAVRMDYETNLSLQSRVHWCRRRPAQFLGYPRT 178
QY 206 ELGMYHLLVQIPIFFTVVVMILTYTKILOALNIRIGTRFSTGQKKARKKKTISLTQ 265
DB 179 SLKRAYAMYQLLLIIFPVLMTSICYARVSAIVYKSKDKRVILSQAMVAFSKAATDAVTF 238
QY 266 HEATDNSQSGGRNVFVGTVTSVVIILARRAVKRHRERR---EROKRVFRMSLLIISTF 322
DB 239 SGYSALPMIITSRN---LKTANTTI-----KSYSNHRRNRAEANKKQIVQMLISIVCMY 290
QY 323 LLCWTPISVLNLTICLG-----PSDLLVKRLCFLVMAYGTTIFHPLLYAFTRQFKQ 376
DB 291 TVCWLP--TIVDELLTSFGYICRTSNTQTCLKMRMGFNALTYCQSCINPILYAFISQNF 349
QY 377 VLK---SKMKRV-VSIVEADPL 395
DB 350 TFKTAYSRMKSRLQGVHSGDPL 372
RESULT 15
A55886
dopamine receptor DIA - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999
C:Accession: A55886
R:Demchislyn, L.L.; Sugamori, K.S.; Lee, F.J.S.; Hamadanizadeh, S.A.; Niznik, H.B.
J. Biol. Chem. 270, 4005-4012, 1995
A:Title: The dopamine D1D receptor. Cloning and characterization of three pharmacolog
A:Reference number: A55886; MUID:95181368; PMID:7876148
A:Accession: A55886
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-450 <DEM>
A:Cross-references: GB:I36877

QY 329 ISVLNTTI-LCLGPSDLLVLRCLFLVMAYGTTFPHLLYAFTRQKQKVLKSKMK 383
Db 361 FFVALVPPFCSSCHMPTLGAIIINWLGYSNLLNPVIYAFNKFQNFQAFKKIIR 416

RESULT 11
T19135
cholecystokinin type A receptor homolog Y54E2A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T19135; T27154
R: Smye, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19078
A:Accession: T19135
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-643 <W1>
A:Cross-references: EMBL:Z81465; PIDN:CA03866.2; GSPDB:GN00020; CESP:Y54E2A.1
A:Experimental source: clone C09F9
R:Lloyd, C.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20319
A:Accession: T27154
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-643 <W12>
A:Cross-references: EMBL:AL032646; PIDN:CAA21687.2; GSPDB:GN00019; CESP:Y54E2A.1
A:Experimental source: clone Y54E2A
C:Genetics:
A:Gene: CESP:Y54E2A.1
A:Map position: 2; 1
A:Introns: 47/3; 71/2; 92/3; 170/2; 251/3; 288/2; 341/3; 428/1; 446/1; 537/2; 573/2; 603/2

Query Match 10.18; Score 222; DB 2; Length 643;
Best Local Similarity 22.2; Pred. No. 3.5e-10;
Matches 94; Conservative 86; Mismatches 188; Indels 56; Gaps 18;

QY 2 CFSPI---LEINQSESNITVRDIDDINTWYQPLSVPLSFQVSLGFLMLEIVL--GL 56
Db 38 CTHPLHKYLEVCI---SRCTVPDD-----TVFFSMTDEELFEIALPGFLYLTFLVGTI 88

QY 57 GSNLTVLVLYCMKSNLINSVNIITMNLHVDLVICVGIPLTIVILLLESNTALICC 116
Db 89 GNSWVIFVWNEFKR--HRNVNIFLASLSTADLCILWFCVP----IMPWYMSHTWSMR 142

QY 117 FHEACVS-----FASVSTAINVFAITLDYDISVKPANRI--LTMGRAVMLMISWIFS 168
Db 143 F--ACYSVHYIQQFTCFCSVLTMWISFERELATAYPMRNWESSIGRAKKVILLINWSS 200

QY 169 FFSFLIPFIEVNF---FSLQSGNT-WENKTLVCVSTNEYYTELGMYYHLVQIPIFFFTV 224
Db 201 AVLAVPTAVRMDYETNLSLSQQRVHWCRRRPPAQLGYPRTSLNKAAYVQLLLIIFPV 260

QY 225 VMLTITTKILQALNIRIGTRFSTGQKKARKKTKTISLTQHEATDMSQSGGRNVFVG 284
Db 261 LTMSCIVARSAIVYKSKDVIISQAWFAESKAATDAVFSGVSAIPMTTSRN----L 316

QY 285 RTSVSVIALLRAVYKRRHRR---ERQKRVFMSLLIISTIFLLCWTPISVLNTTILCIG- 340
Db 317 KTANTTI-----KSYSNHRNNRVAEANKQIVQMLISIVCMYTCWLP-TIVDELLTSEGY 371

QY 341 -----PSDLLVLRCLFLVMAYGTTFPHLLYAFTRQKQKVLK---SKMKRV-VSIVE 391
Db 372 ICRTSNTQTLKMRMGFNALTYCOSCINPILYAFISONFRSTFTKAYSRMKSLQGVGHS 431

QY 392 ADPL 395
|||
Db 432 GDPL 435

RESULT 12

148231
serotonin receptor 5B - mouse
N:Alternate names: 5-hydroxytryptamine 5B receptor (5HTR-5c)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 148231
R:Matthes, H.; Boschart, U.; Amlaiki, N.; Grailhe, R.; Plassat, J.L.; Muscatelli, F.; Mol. Pharmacol. 43, 313-319, 1993
A:Title: Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors define a new
A:Reference number: 148231; MUID:93196607; PMID:8450829
A:Accession: 148231
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-370 <RES>
A:Cross-references: EMBL:X69867; NID:G288735; PIDN:CAA49501.1; PID:G288736
C:Superfamily: octopamine receptor type I
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; tran

Query Match 10.08; Score 218.5; DB 2; Length 370;
Best Local Similarity 23.48; Pred. No. 3.7e-10;
Matches 89; Conservative 56; Mismatches 141; Indels 95; Gaps 12;

QY 39 SFQVSLTGFLMLEIVLGLGSLNLTVLVLYCMKSNLINSVNIITMNLHVDLVICVGIPL 98
Db 50 AFTVLVTVLLVLLIAATFLMNLVLT-ILVRAPHRYPHNLVASTAYSDVLVAVLVNPL 108

QY 99 TIVILL-----LSLENTALICCPHEACVSPASVSTAINVFAITLDY- 141
Db 109 SLVSEKAGRRWQLGRSLRCHWISFD---VLCC-----TASITWVAALADRYW 154

QY 142 DISVKPANRIITMGRAVLMISI-WIFFSFLPIPIEVNFESLQSGTWNKTLICVST 200
Db 155 TITRHLQVTLTRGRASALMAITWALSALIALAPLL-----FGWGEAYDARLQRCQVS 208

QY 201 NEYTELGMYYHLVQIPIFFFTVVMVLTITTKILQALNIRIGTRF-----STGQKKK 253
Db 209 QE-----PSYAVESTCGAFVLPVAVLVFYWKYKAAKFRGRRRAVPLPATTOAKE 262

QY 254 ARKKKTISLTQHEATDMSQSGGRNVFGRVTSVVIALLRAVYKRRERERQKRVFR 313
Db 263 APPESEMFTARRATVTFQTS-----DSWREQEKKRAAM 298

QY 314 MSLIISITFLCWTPISVLNTITLCLGPSDLLVLRCLG-----FLVMAYGTTFHP 364
Db 299 MVLIGLVFVLCWIPFEL-----TELISPLCACSLPPIWKSIFLWLGYSNFFNP 348

QY 365 LLYAFTROKQKQKVLKSKMKR 385
Db 349 LIYTAFAKNYNNNAFKSLFTKQ 369

RESULT 13
A47430
gastrin/cholecystokinin receptor B, short splice form - human
N:Alternate names: CCK-B/gastrin receptor; cholecystokinin-B/gastrin receptor
C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: A47430; JCI352; A46645; A48262
R:Ito, M.; Matsui, T.; Taniguchi, T.; Tsukamoto, T.; Murayama, T.; Arima, N.; Nakata, J. Biol. Chem. 268, 18300-18305, 1993
A:Title: Functional characterization of a human brain cholecystokinin-B receptor. A t
A:Reference number: A47430; MUID:93352657; PMID:8349705
A:Accession: A47430
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-447 <ITO>
A:Cross-references: GB:D13305; NID:9436039; PIDN:BAA02564.1; PID:9436040
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:136448)
R:Pisegna, J.R.; de Weerth, A.; Huppi, K.; Wank, S.A. Biochem. Biophys. Res. Commun. 189, 296-303, 1992
A:Title: Molecular cloning of the human brain and gastric cholecystokinin receptor: S
A:Reference number: JCI352; MUID:93080572; PMID:1280419

QY 357 YGTTTFPHLLYAFTRQKFKVLSKM-----KKRVSVIVEADPLPNNAVHNSWID 407
Db 309 MTSIVNPLYVFMKSRHRRALKDDMTLTNARRHTNVGLSRFTSPSVSVYVRTLE 367

RESULT 9
A46195
cholecystokinin B receptor subtype - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Apr-2000
C:Accession: A46195
R:Wank, S.A.; Pisegna, J.R.; de Weerth, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 8691-8695, 1992
A:Title: Brain and gastrointestinal cholecystokinin receptor family: structure and function
A:Reference number: A46195; MUID:92409582; PMID:1528881
A:Accession: A46195
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-452 <WAK>
A:Cross-references: GB:M99418; NID:g203459; PIDN:AAA40925.1; PID:g203460
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIN:114083, NCBI:P:114084)
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 10.2%; Score 222.5; DB 2; Length 452;
Best Local Similarity 22.9%; Pred. No. 2.2e-10;
Matches 87; Conservative 73; Mismatches 145; Indels 75; Gaps 12;

QY 36 YPLSFQVSLTGFLMLEIVLGLGSLNLTVLVLYCMKSNLINSVNIITMNLHVLVDVLCVGC 95
Db 61 YAVIFLMSVGGNVLITVLGL-----SRLRTVTNFAFLLSLAVSDLLAVAC 107

QY 96 IPLTIVILLLSLESNTALICCFHEACVSF-----ASVSTAINVEATLDRYDISVKP--A 148
Db 108 MFETLPLNMGVTFGTVC-----KAISYLMGVSVSVST-LNLVALERYSIAICPLQA 162

QY 149 NRILTMGRVAMLMISWIFSFSLIPFTIEVNFSS-----LQSGTNWKNKTLILCVSTNE 202
Db 163 RVWQTRSHAARVILATWLLSLG-LMVPYPVYTMQVPGPRVLQCMHRWFS-----211

QY 203 YTELGMYHLLVQIDIFFETVVVMLITYTKILQALNIRIGTRFSTGQKKARKKKTISL 262
Db 212 --ARVQQTWSVLLLLFPFGVIAVAYGLISRELYLGLHFHDGENDSETQSRARNOGGL 269

QY 263 TQHEATMSQSGGRNVFVGRTSVSVIILRRVAKRHRERRE-----306
Db 270 PGGAAPCPVHQGGRPV-----TSVAGEDSGCCVQLPRSRLEMTTLTPPGVPVGPGR 324

QY 307 -----RQKRVFRMSLLIISTFLLCWTPISVNLITILCLGP--SDLLVKLRLCFL-VMA 356
Db 325 PNOAKLLAKKRVYRMLLVILVFLLCWLPVSVNTWRAFDGPGAQRLSGAPISFIHLLS 384

QY 357 YGTTTFPHLLYAFTRQKFK 376
Db 385 YVSACVNPLYVCFMHRFRFQ 404

RESULT 10
138209
serotonin receptor 1A - human
N:Alternate names: 5-hydroxytryptamine receptor 1A (5-HT1A)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 15-May-1998 #text_change 21-Jul-2000
C:Accession: I38209; I37104; S07343; I56176; S31438
R:Jacobsen, J.S.; Refolo, L.M.; Conley, M.P.; Sambamurti, K.; Humayun, M.Z.
Mutat. Res. 179, 89-101, 1987
A:Title: DNA replication-blocking properties of adducts formed by aflatoxin B1-2,3-dichloroacetate
A:Reference number: I38209; MUID:87258013; PMID:3110609
A:Accession: I38209
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-422 <JAC>

A:Cross-references: EMBL:X57829; NID:g36428; PIDN:CAA40962.1; PID:g36429
R:Parks, C.L.; Chang, L.S.; Shenk, T.
Nucleic Acids Res. 19, 7155-7160, 1991
A:Title: A polymerase chain reaction mediated by a single primer: cloning of genomic DNA
A:Reference number: I37104; MUID:92115564; PMID:1766875
A:Accession: I37104
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <PAR>
A:Cross-references: EMBL:Z11168; NID:g1033027; PIDN:CAA77560.1; PID:g1033028
R:Kobilka, B.K.; Frielle, T.; Collins, S.; Yang-Feng, T.; Kobilka, T.S.; Francke, U.;
Nature 329, 75-79, 1987
A:Title: An intronless gene encoding a potential member of the family of receptors co-
operates with a G protein-coupled receptor in the regulation of cyclin D1 expression
A:Reference number: S07343; MUID:87315369; PMID:3041227
A:Accession: S07343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151, 'PR', 155-171, 'I', 173-417, 'N', 419-422 <KOB>
A:Cross-references: EMBL:X13556; NID:g35523; PIDN:CAA31908.1; PID:g35524
A:Note: the authors translated the codon ATC for residue 172 as Met
R:Aune, T.M.; McGrath, K.M.; Sarr, T.; Bombara, M.P.; Kelley, K.A.
J. Immunol. 151, 1175-1183, 1993
A:Title: Expression of 5HT1a receptors on activated human T cells. Regulation of cyclin
D1 expression
A:Reference number: I56176; MUID:93329096; PMID:8393041
A:Accession: I56176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'RPR', 203-227, 'R', 229-244, 'A', 245-354, 'T', 356-362, 'MRP', <NUN>
A:Cross-references: GB:S64045; NID:9404416; PIDN:AAD13945.1; PID:94261645
C:Genetics:
A:Gene: GDB:HTR1A; ADRB2RL1; ADRBRL1
A:Cross-references: GDB:120686; OMIM:109760
A:Map position: 5cen-5q11
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; neurotransmitter r
eceptor
F:37-62/Domain: transmembrane #status predicted <TM1>
F:74-98/Domain: transmembrane #status predicted <TM2>
F:110-132/Domain: transmembrane #status predicted <TM3>
F:153-178/Domain: transmembrane #status predicted <TM4>
F:192-217/Domain: transmembrane #status predicted <TM5>
F:346-367/Domain: transmembrane #status predicted <TM6>
F:379-403/Domain: transmembrane #status predicted <TM7>
F:10,11,24/Binding site: carboxylate (Asn) (covalent) #status predicted
F:109-187/Disulfide bonds: #status predicted
F:420/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 10.1%; Score 222; DB 2; Length 422;
Best Local Similarity 21.9%; Pred. No. 2.2e-10;
Matches 91; Conservative 70; Mismatches 153; Indels 102; Gaps 15;

QY 38 LSPQV---SLTGFLMLEIVLGLGSLNLTVLVLYCMKSNLINSVNIITMNLHVLVDVLCVGC 94
Db 33 VSVQVITSLLLGLTFLICAVLG---NACVNAIALESRLQN-VANYLIGSLAVTDLWVSVL 88

QY 95 CIPLTIVILLLSLESNTALICCFHEACVSFASVSPASVAINFAITLDRYDISVKPANKR--L 152
Db 89 VLPMAALYQVNLKWTGLGQVTCDLFIALDVLCCTSSILHLCAIALDRYWAITDPIDYVNR 148

QY 153 TMGRVAMLMISWIFSFSLIPFTIEVNFSLQSGTNWKNKTLILCVSTNEY-----YTE 206
Db 149 TPRRAAALISLTWILGLISIPPM-----GWRTPEDRSDPDACITSKDHGYTIYST 200

QY 207 LGMYHLLVQIDIFFETVVVMLITYTKILQALNIRIGTRFSTGQKKARKKKTISLTQH 266
Db 201 FGAFY-----IPL-----LLMVLVYGRIFPAARFRI-----RKTVKVKEVTGADTRH 242

QY 267 EATDMSQ-----SSGGRNVFVGRTSVSVIILRRVAK-----299
Db 243 GASPAQPKKSVNGESGRNWRVGESKAGCALCANGAVRQGDGAALIEVIEVHRVGNK 302

QY 300 -----RHRERERQKRVFRMSL-----330
Db 303 EHLPLPSEAGPTPCASPASFERKNERNAEKR--KMALARERKTVKTLGIIMGTFILCWLP 360

[illegible]

F:57-82/Domain: transmembrane #status predicted <TM1>
F:93-119/Domain: transmembrane #status predicted <TM2>
F:131-151/Domain: transmembrane #status predicted <TM3>
F:173-193/Domain: transmembrane #status predicted <TM4>
F:225-249/Domain: transmembrane #status predicted <TM5>
F:330-348/Domain: transmembrane #status predicted <TM6>
F:366-389/Domain: transmembrane #status predicted <TM7>

Query Match 12.2%; Score 266.5; DB 2; Length 444;
Best Local Similarity 23.8%; Pred. No. 5.8e-14;
Matches 102; Conservative 85; Mismatches 131; Indels 111; Gaps 20;

QY 6 ILEINMQSNTIVRDDIDDINTNMV-----QP-LSYPLSFQVSLTGFLMLEIVLGLGNS 59
DB 18 VVDSLMLNGSNITPPCELGLENTEFLCDLPQPSKEWQPAVLIIYSLFLSVLGLNTLV 77
QY 60 LTVLVLYCMKSNLINSVNIITMNLHVLDVIVCGCIPITIVILLLSLESNTALICCFHE 119
DB 78 IIVLI-----RNKRMRTVTNIFLLSLAVSDMLCLFCMPENLIPNLKDFIFGSAVC---K 130
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRAVLMIMISWIFSF----- 169
DB 131 TTTYFMGTSVSVSTENLVAISLERYGAICRPLQSRVWQTKSHALKVIAATWCLSFMTMP 190
QY 170 ---FSLPIEFVNV-----FSLQSG---NWNKNTLLCVSTNEYYTELGMYYHLLVQ 216
DB 191 YPIYSNLVFPFTKNNOTANNCREFLPSDAMQOSWQTFLLI----- 231
QY 217 IPIFFTVVMLTYTKILOALNIRIGTRFSTGQKKARKKK-TISLTTQHEATD----- 270
DB 232 --LFLPLGVVWVAYGLI--SLEYQGIKFDASQKSAKEKPPSTGSGTRYEDSDGCIYQ 287
QY 271 -----MSQSSGGRNVVFGVRTSVSV--IIALRAVRKRRREROKRVRMSL 316
DB 288 KSRRPKLELQQLSSGSGSR-LNRRSSSSAANLTA-----KKRVIRMLI 332
QY 317 LIISTFELLQWTPISVLT-----TI-----LCIGPSDLLVKRLCLFLVMAYGTTIFHP 366
DB 333 VIVVLEFLCWMPITFSANAWRAVDTVSAEKHLGTPISFI-----LLLSYTSSCVNP 385
QY 367 YAFTRQKFQ 375
DB 386 YCFMNRFR 394

RESULT 5
JN0692
cholecystokinin type A receptor - human
C:Species: Homo sapiens (man)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Apr-2000
C:Accession: JN0692; JN0590
R:de Weerth, A.; Pilegna, J.R.; Huppi, K.; Wank, S.A.
Biochem. Biophys. Res. Commun. 194, 811-818, 1993
A:Title: Molecular cloning, functional expression and chromosomal localization of the hu
A:Reference number: JN0692; MUID:93343941; PMID:8343165
A:Accession: JN0692
A:Molecule type: mRNA
A:Residues: 1-428 <DEW>
A:Cross-references: GB:L19315; NID:q306595; PIDN:AAA02819.1; PID:q306596
A:Experimental source: gallbladder
R:Ulrich, C.D.; Ferber, I.; Hollicky, E.; Hada, E.; Buell, G.; Miller, L.J.
Biochem. Biophys. Res. Commun. 193, 204-211, 1993
A:Title: Molecular cloning and functional expression of the human gallbladder cholecysto
A:Reference number: JN0590; MUID:93277552; PMID:8503909
A:Accession: JN0590
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-428 <ULR>
A:Cross-references: GB:L13605; NID:q306490; PIDN:AAA35659.1; PID:q306491
A:Experimental source: gallbladder
C:Comment: This protein has diverse physiological roles in the gastrointestinal system w
ch, and secretion from gastric mucosal cells.
C:Genetics:

A:Gene: GDB:CKAR
A:Cross-references: GDB:141927; OMIM:118444
A:Map position: 4pter-4qter
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotei
F:40-67/Domain: transmembrane #status predicted <TM1>
F:78-104/Domain: transmembrane #status predicted <TM2>
F:116-137/Domain: transmembrane #status predicted <TM3>
F:158-178/Domain: transmembrane #status predicted <TM4>
F:208-234/Domain: transmembrane #status predicted <TM5>
F:314-332/Domain: transmembrane #status predicted <TM6>
F:350-369/Domain: transmembrane #status predicted <TM7>
F:10,24,190,299/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:10,24,190,299/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #sta
F:256/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predi

Query Match 11.5%; Score 253; DB 2; Length 428;
Best Local Similarity 22.8%; Pred. No. 6.9e-13;
Matches 95; Conservative 85; Mismatches 132; Indels 104; Gaps 18;

QY 15 SNITVRDDIDDINTNMV-----OPLSVPLSFQVSLTGFLMLEIVLGLGNS 59
DB 12 SNITPPCELGLENTEFLCDLPQPSKEWQPAVLIIYSLFLSV-----LGNT 60
QY 60 LTVLVLYCMKSNLINSVNIITMNLHVLDVIVCGCIPITIVILLLSLESNTALICCFHE 119
DB 61 LVITVL--TRNKRMTVTNIFLLSLAVSDMLCLFCMPENLIPNLKDFIFGSAVC---K 115
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRAVLMIMISWIFSF----- 169
DB 116 TTTYFMGTSVSVSTENLVAISLERYGAICRPLQSRVWQTKSHALKVIAATWCLSFMTMP 175
QY 170 ---FSLPIEFVNVFSLQSGNTWENKTL-LC--VSTNEYYTELGMYYHLLVQIPFFFT 223
DB 176 YPIYSNLVFPFTKNN-----NQTANMCRFLPNV---MQOSWTFLLIILFLIP 221
QY 224 VVMLTYTKILOALNIRIGTRFSTGQKKARKKKTISLTT-QHEATD----- 270
DB 222 GIWMVWVAYGLI--SLEYQGIKFEASQKSAKEKPPSTGSGTRYEDSDGCIYQAKTRPRK 279
QY 271 -----MSQSSGGRNVVFGVRTSVSVIILARRVRREROKRVRMSLIIISTFLLC 325
DB 280 LELRQLSTGSSSRANRIRNSAANLMA-----KKRVIRMLIVIVLFLFLC 325
QY 326 WTPISVLTNTILCLGPSDLLVKRLC-----FLVMAYGTTIFHPLLYAFTRQKFQ 375
DB 326 WMPIFSANA--WRAYDTASAERRLSGTPISFILLISYTSSCVNPITYCFMNRFR 378

RESULT 6
JQ1614
gastrin receptor - multimammate rat (Mastomys natalensis)
C:Species: Mastomys natalensis
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: JQ1614
R:Nakata, H.; Matsui, T.; Ito, M.; Taniguchi, T.; Naribayashi, Y.; Arima, N.; Nakamur
Biochem. Biophys. Res. Commun. 187, 1151-1157, 1992
A:Title: Cloning and characterization of gastrin receptor from ECL carcinoma tumor of
A:Reference number: JQ1614; MUID:92412082; PMID:1530611
A:Accession: JQ1614
A:Molecule type: mRNA
A:Residues: 1-450 <NAK>
A:Cross-references: GB:D12817; NID:q220646; PIDN:BAA02250.1; PID:q220647
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane
F:59-83/Domain: transmembrane #status predicted <TM1>
F:87-109/Domain: transmembrane #status predicted <TM2>
F:132-150/Domain: transmembrane #status predicted <TM3>
F:172-188/Domain: transmembrane #status predicted <TM4>
F:216-243/Domain: transmembrane #status predicted <TM5>
F:334-357/Domain: transmembrane #status predicted <TM6>
F:380-398/Domain: transmembrane #status predicted <TM7>
F:7,30,36/Binding site: carbohydrate (Asn) (covalent) #status predicted

[illegible][illegible]

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|-----|--------|--------|----|---------------------|
| | | Match | % | | | | |
| 1 | 282 | 12.9 | 430 | 2 | I51898 | | cholecystokinin A |
| 2 | 268 | 12.2 | 427 | 2 | S50150 | | gastric CCK-A rece |
| 3 | 287 | 12.2 | 436 | 2 | JC5599 | | cholecystokinin-A |
| 4 | 266.5 | 12.2 | 444 | 2 | A42685 | | cholecystokinin re |
| 5 | 253 | 11.5 | 428 | 2 | JN0692 | | cholecystokinin ty |
| 6 | 237.5 | 10.8 | 450 | 2 | QJ1614 | | gastrin receptor - |
| 7 | 229.5 | 10.5 | 584 | 2 | JC7809 | | sulfakinin recepto |
| 8 | 224 | 10.2 | 539 | 2 | T27559 | | hypothetical prote |
| 9 | 222.5 | 10.2 | 452 | 2 | A46195 | | cholecystokinin B |
| 10 | 222 | 10.1 | 422 | 2 | I38209 | | serotonin receptor |
| 11 | 222 | 10.1 | 643 | 2 | T19135 | | cholecystokinin ty |
| 12 | 218.5 | 10.0 | 370 | 2 | I49231 | | serotonin receptor |
| 13 | 218 | 9.9 | 447 | 2 | A47430 | | gastrin/cholecysto |
| 14 | 217 | 9.9 | 580 | 2 | H87963 | | protein Y54E2A.1 [|
| 15 | 216 | 9.9 | 450 | 2 | A55886 | | dopamine receptor |
| 16 | 214 | 9.8 | 369 | 2 | S38744 | | serotonin receptor |
| 17 | 213.5 | 9.7 | 357 | 2 | B47472 | | 5-hydroxytryptamin |
| 18 | 213.5 | 9.7 | 452 | 2 | JC2459 | | gastrin/cholecysto |
| 19 | 213.5 | 9.7 | 509 | 2 | A47174 | | serotonin receptor |
| 20 | 213 | 9.7 | 453 | 2 | S32817 | | gastrin receptor - |
| 21 | 210.5 | 9.6 | 357 | 2 | S28058 | | serotonin receptor |
| 22 | 209.5 | 9.6 | 357 | 2 | I37107 | | 5-HT5A serotonin r |
| 23 | 209.5 | 9.6 | 409 | 2 | I48095 | | A2 adenosine recep |
| 24 | 209.5 | 9.6 | 455 | 2 | T15622 | | hypothetical prote |
| 25 | 208 | 9.5 | 515 | 2 | A40491 | | alpha-1-adrenergic |
| 26 | 207.5 | 9.5 | 517 | 2 | A45121 | | alpha-1B adrenergic |
| 27 | 206.5 | 9.4 | 326 | 2 | A38144 | | adenosine receptor |
| 28 | 206 | 9.4 | 421 | 2 | I49375 | | serotonin receptor |
| 29 | 204 | 9.3 | 484 | 2 | A45490 | | neuropeptide Y/pep |

This Page Blank (uspto)

```
; TYPE: PRT
; ORGANISM: Human
US-10-044-592-12

Query Match      8.8%; Score 193.5; DB 12; Length 370;
Best Local Similarity 20.9%; Pred.No.2.9e-10;
Matches 76; Conservative 66; Mismatches 119; Indels 103; Gaps 14;

QY   44 LTGFLML-----EIVLGLGSNLTVLVLYCMKSNLSNSNIITMNLHVLDVIICVGCIPLT 99
    | :| :|| :||| :||| : :||| : || :||| :|||
Db   59 LKGLIIVLLSVVVVGLVGN-CLLVLIARVRRLHNVTFNLGNLALSDVLMCTACVPLT 117

QY   100 IVI-----LLLLLESNTALICCPCHEACVSFASVSTAINVFATILDRYDIS 144
    | :| :|| :||| :||| : :||| : || :||| :|||
Db   118 LAYAFEPGRGVFGGLCHLVFFLPQPTVTVYSVP-----TLTTIAVDRIYVL 163

QY   145 VKPANRILTGRAVMIMISWIFSFSLPIEFVNFFSLOSGNTWENKTLLCV---STN 201
    | :| :|| :||| :||| :||| :||| :||| :|||
Db   164 VHPLRRIISRLSAVAVALWALSAVLAPAAVHTVHVLEK-----PHDVRLEEWFWSQ 218

QY   202 EYTTELGMYYHLVQIPIFFTVMVMLITYTKILQALNIRIGTRPSTGOKKKARKKKTIS 261
    | :| :|| :||| :||| :||| :||| :||| :|||
Db   219 ERQRQLYAWGLLV---TYLLPLLVLILLSYRV-----SVKLNRNVPG 259

QY   262 LTTOHEATDMQSOGSRNNVFGVRTSVIIALLRAVKRRHRERROKRVRFRMSLLIIST 321
    | :| :|| :||| :||| :||| :||| :||| :|||
Db   260 CVTQSOA-DWD-----RARRRTFCLLVVVVVV 286

QY   322 FLWCWPISVLNNTIILCLGPSDL-----LVKLRLCFLVMAYGTTFIFHPLLYAFTQKFK 375
    | :| :|| :||| :||| :||| :||| :||| :|||
Db   287 FAVCWLPPLHVEN-LLRDLDPAIDPYAGLVQL-LCHW-LAMSSACYNPFIYAWLHDSFR 343

QY   376 KVLK 379
    :| :
Db   344 EEEL 347

Search completed: December 4, 2002, 16:03:31
Job time : 14 secs
```


Db 277 SPEKVLMDGSRKDKALPNSGDETLM--RRST-----IGKSVQTSNEORASKVL 326
QY 316 LLIISTELLCTPISVLN--TTILCLGPSDLLVLR--CFLVMAYGTFIFHPLLYAFTRQK 373
Db 327 GIVFFLLMWCFFITNITLVLCDSNQTTLQMLLEIFWVGVSQVPLVYTFLENKT 386
QY 374 FQ-----KVLKS--KMKKRVYSIVEADPLPNAV-----IHNS-----WIDPK 409
Db 387 FRDAGRYITCNRYATSKVTLRRSSKIYFRNPMAENSFFKKHGIRNGINPAMYQSPM 446
QY 410 RNKKTFEEDSEI 421
Db 447 RLRSSTIQSSSI 458
RESULT 12
US-09-919-497-72
; Sequence 72, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-72

Query Match 8.9%; Score 195; DB 10; Length 481;
Best Local Similarity 20.8%; Pred. No. 3e-10;
Matches 90; Conservative 83; Mismatches 167; Indels 92; Gaps 19;

QY 48 LMLEIVGLGSNLTVLVLYCHKSLNLSVNIIPMNLHVLVDVLCGICPLTIVILLLSL 107
Db 61 ILMWIIPITGNTLVILAVLSLEKKL-QYATNYFLMSLAVADLLVGLFVMPALTITMFEA 119
QY 108 ESNTALICC---PHEACVSFASVSTAINVFAITLDRYDISVKP--ANRIITMGRVYLM 161
Db 120 MWPLVLCFANFLDLVLFSTASI---MHLCAISVDVYIAIKKPIQANQYNSRATAFIKI 176
QY 162 ISIWIFSFSLPIEFIEVNFPSLOSNGTWENKTLTLCVSTNEYTELGMYYHLVQIPIFF 221
Db 177 TVVWLIS-IGTAIP-VPIKIGLTDVDP--NNITCVLTKEREGD---FMLFGSLAAPP 227
QY 222 FTVVVMLITTKILOALNIRIGTRFSTGQKKKARKK-----TISLTTQHEATDMS 272
Db 228 TPLAIMIVTYFLTHAL-----QKAVLVKNKPPQRLTWLTVSTVVFQDETPCS 276
QY 273 Q-----SSGRRNVFGVTSVSVIILALRRVAKRHRERQKRVFRMS 315
Db 277 SPEKVLMDGSRKDKALPNSGDETLM--RRST-----IGKSVQTSNEORASKVL 326
QY 316 LLIISTELLCTPISVLN--TTILCLGPSDLLVLR--CFLVMAYGTFIFHPLLYAFTRQK 373
Db 327 GIVFFLLMWCFFITNITLVLCDSNQTTLQMLLEIFWVGVSQVPLVYTFLENKT 386
QY 374 FQ-----KVLKS--KMKKRVYSIVEADPLPNAV-----IHNS-----WIDPK 409
Db 387 FRDAGRYITCNRYATSKVTLRRSSKIYFRNPMAENSFFKKHGIRNGINPAMYQSPM 446
QY 410 RNKKTFEEDSEI 421
Db 447 RLRSSTIQSSSI 458
RESULT 13

US-09-804-551B-8
; Sequence 8, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804,551B
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-8

Query Match 8.9%; Score 195; DB 10; Length 518;
Best Local Similarity 21.2%; Pred. No. 3.3e-10;
Matches 81; Conservative 66; Mismatches 129; Indels 106; Gaps 13;
QY 53 VLGGSNLTVLVLY-CMKSNLINSVNIITMNLHVLVDVLCGICPLTIVILLLSLSENT 111
Db 75 VLVFGN--VLVCYVLRNRAMQVITNITNLSALSDLLCLVAVPPTPLTYFMGRWAFG 132
QY 112 ALIC---CFHEACVSFASVSTAINVFAITLDRYDISVKPANRIITMGRVYLMISWIFS 168
Db 133 RSLCHLVSAAGCSIIYSTLT---LTSIAIDRYEVIYFPHPMKLSTCIGIIVSIWIA 189
QY 169 FFSFLPIEFIEVNFPSLOSNGTWENKTLTLCVSTNEYTELGMYYHLVQIPIFFFTVVVML 228
Db 190 LLA-TVPEY-----GMYMKTNELVNGTGTGNETL 217
QY 229 ITYTKILOALNIRIGTRFSTGQKKKARKKKTISLTTQHEATDMSQSGRRNV----- 280
Db 218 VEATLMLNGSFVAGSGFIEAPDSTS-----ATQAYMQVMTAGSTGPEMPYRVYCE 269
QY 281 -----VFGVTS-----VSVIILR-----RAVKRHRER 306
Db 270 ENWPFSEQYRVFGAITTLPVLPFFIISICYVWISVKLNQARAKPCKSSRREEDRD 329
QY 307 RQKRVFMSLLIISTFLCTWTPISVLNNTILCLGPSDLLVKLR-----LCFLV---MAY 357
Db 330 RKRTRNMLIAMVAVFGLSWLPINNVNIF-----DDFDDKSNWRFFYILFFVFAHSAM 383
QY 358 GTTIFHPLLYAFTRQKFKVLK 379
Db 384 SSTCYNFFLIANLENFRKEFK 405

RESULT 14
US-10-044-592-12
; Sequence 12, Application US/10044592
; Patent No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: polypeptides, their production and use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/Jp98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 12
; LENGTH: 370

Db 3 DYNSTINSLSTRVLAFFMSLVAP---AIMLG---NALVILAFVVDKNLRHR--SSYPFL 55
Qy 83 NLHVLVDVIVCGICPILVILLLESNTALICCFCHEACVSPASVSTAINVFAITLDRYD 142
Db 56 NLAISDFGVGISIPLYPHILFEWDFGE-ICVFWLITDYLCCASVYNIVLISYDRY- 113
Qy 143 ISVKPANRILTMGRVAVM-----LMISWIFSFSLPIPIEYVNFSSQSGTWNKTLICV 198
Db 114 LSVNAVSVYRTHGTGVLKTVTLWAVVLAFL-----VNGPMILVSESKDEGSEC- 164
Qy 199 SINEYTTLMGY--YHLLVQIPDIPFTVVVMLITTKI-----LQ 236
Db 165 -----EPGFSEYIILAITSFLEVPVILVAVFNMNIVWSLWKROHLSRCQSHPGLT 217
Qy 237 ALNIRI-GTRFSTGOKKARKKTTISLTTOHEATDMSOSSGGR-NVFGVRTSV-SVITA 293
Db 218 AVSSNICGHSF-----RGLSSRRSLSASTEVPASHSRQRKSSLMFSRRKMSNTIA 273
Qy 294 LRR-----AVKRH-RERR--RQKRVFRMSLLIISTFLLCWTPIIS-----VLNTTILCL 339
Db 274 SKMGFSQSQSDVALHOREHVELLRARLAKSLAILLGVFAVCWAPVSLFTIVLSFYSSAT 333
Qy 340 GPSDLLVLRCLFLVMAVGTTFHPLLYAFTRQKQVLRKSKMKRVVSIIVEADPLPN 397
Db 334 GPKS--VMYRIAFWLQWNSFV-NPLLYPLCHKRFQKAFK-----IFCIKKQPLPS 382
RESULT 10
US-10-005-010-4
; Sequence 4, Application US/10005010
; Patent No. US20020115149A1
; GENERAL INFORMATION:
; APPLICANT: Weinshank et al, Richard L.
; TITLE OF INVENTION: Method of Obtaining A Composition Comprising A 5-HT1D
; TITLE OF INVENTION: Selective Compound
; FILE REFERENCE: 36536-B4
; CURRENT APPLICATION NUMBER: US/10/005,010
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/371,705
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-005-010-4
Query Match 8.9%; Score 195; DB 12; Length 390;
Best Local Similarity 19.5%; Pred. No. 2.3e-10;
Matches 78; Conservative 90; Mismatches 175; Indels 58; Gaps 14;
Qy 5 PILENMOSESNIYRDDIDINTWYQPLSPFQVSLTGFL-MLEIVLGLGSLNVLVL 63
Db 21 POANLESAPSONCSARDYI-----YQD-SISLPKVLVLLALITLITLSNAFVIA 72
Qy 64 VLYCKMSNLINSVSNITMNLHVLVDIVCGICPILVILLLESNTALICCFCHEACVS 123
Db 73 TVI--RTRKLHTPANLIASLAVTOLLVSILVMPITMTVTVGRWTLGOVVCDFWLSSDI 130
Qy 124 FASVSTAINVFAITLDRY-----DISVKPANRILTMGRVAVLMISWIFSFSLPIPIEY 179
Db 131 TCCITASILHLCVIALDRYWAITDAVEYSKR--TPKRAVMIALVWFSISISLPPFF-- 186
Qy 180 NFPSLQSGNTW-----ENKTLICVSTNEYTELGMYYHLLVQIPDIPFTVVVMLITY---- 231
Db 187 -----WRQAAEEVSECVVNTDTH-----ILYTVYSTVGAFYFPTLLILYALYRI 231
Qy 232 -----TKILQALNIRIGTRFSTGOKKARKKTKTISLT--QHEATDMSQSSGGRNVFVGR 285
Db 232 VYEARSRLIKQTPNRTGKRLTRAQLITDPSGTSSTSVTSINSRVPDVPSESGSPVTVNQVK 291

Qy 286 TSVSVIIARRAVKRHRERRRQKRVFRMSLLIISTFLLCWTPIISVNLNTI-LCLGPSDL 344
Db 292 VYSDALLEKKLMAER-----KATKTLGITLGAIVCWLPFFIISLVMPICKDACWF 346
Qy 345 LVYLRCLFLVMAVGTTFHPLLYAFY-----RQKQVLRKSK 381
Db 347 HLAIFDFFTWLGLNSLINPIIYITMSNEDFKQAFHKLIRFK 387
RESULT 11
US-08-681-219-31
; Sequence 31, Application US/08681219
; Patent No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,219
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-681-219-31
Query Match 8.9%; Score 195; DB 8; Length 481;
Best Local Similarity 20.8%; Pred. No. 3e-10;
Matches 90; Conservative 83; Mismatches 167; Indels 92; Gaps 19;
Qy 48 LMLEIVLGLGSLNLTVLVLYCMKSNLSVSNITMNLHVLVDIVCGICPILFTIVILLSL 107
Db 61 ILMWIPIITGGNTVLIVAVLSLEKL-QYAFNYFLMSLAVADLLVGLFVMPIALTIMFER 119
Qy 108 ESNTALICC-----FHBACVSVASVSTAINVFAITLDRYDISVKP--ANRILTMGRVAVLM 161
Db 120 MWPLPLVLCPAWFLDLVLFSTASI---MHLCAISVDRYIAIKKPIQAOVNSRATAFIKI 176
Qy 162 ISWIFSFSLPIPIEYVNFSSQSGTWNKTLICVSTNEYTELGMYYHLLVQIPDIP 221
Db 177 TVVWLIS-TGIAIP-VPIKGIETVDNPN---NNITCVLTKERFGD----FMLFGSLAAFP 227
Qy 222 FTWVWMLIYTILOALNIRIGTRFSTGOKKARKKTK-----TISLTTOHEATDMS 272
Db 228 TPLAINIVTYFTLIAL-----QKAYLVKNKPPQRLTWLTVTSTVQFQDPTPCS 276
Qy 273 Q-----SSGGRNVVFGVRTSVSVIIARRAVKRHRERRRQKRVFRMS 315


```

Db 51 IILGVGNL-ALIIILKQKEMRNVTNIIIVNLSFSDLLVAIMCLPFTFYVYTLMDHWVFG 101
QY 107 ---LESNTALICCFHEACVSFASVSTAINFAITL---DRYDISV-----KPNRILTMG 155
Db 110 EMCKLNPFFVQC-----VSITVSFSLVLAVERHQLIINPRGWRPNRRHAYVG 158
QY 156 RAVMLMISIWIFSFPSFLPIPIEVNPFSLQSGNTWENKTL-----LCVSTNEYYTEL-GM 209
Db 159 IAV-----IWLAVASSL-PFL---IYQVMTDEPFQNVTLDAKYDKVYCFDQFPSPDSHRL 209
QY 210 YYHLLVQIPIFFFTVVVYMLITYTKILQALNIRIGTRFSTGQKKKARKKKTISLTQHEAT 269
Db 210 SYTTLLLVLYQFGLPCFIFICYFYI-----IRLKRNNMMDK-----247
QY 270 DMSOSSGRNVVGVRTSVSVIITALRAVKRHRERERQKRVFRMSILLIISTFLLCWTPI 329
Db 248 -----MDNRYRSETKRINIMLLSIVVAFVACWLPL 279
QY 330 SVLNTT-----IILGPSDLLVKRLCLFVLVMAYGTTIFPHLLYAFTRQKFQKVLK 379
Db 280 TIFNTVFDWNHQIATCNHLL--FLCHLTAMISTCV-NPIFYGLNKNFQRLDQ 332

RESULT 6
US-09-771-956-3
; Sequence 3, Application US/09771956
; Patent No. US2001003147A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Michele
; APPLICANT: Brockbeck, Robbin
; APPLICANT: Krause, James
; TITLE OF INVENTION: Chimeric Neuropeptide Y Receptors
; FILE REFERENCE: N2000.001
; CURRENT APPLICATION NUMBER: US/09/771,956
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-771-956-3

Query Match 9.3%; Score 203; DB 10; Length 382;
Best Local Similarity 21.2%; Pred. No. 4.2e-11;
Matches 89; Conservative 72; Mismatches 120; Indels 138; Gaps

QY 52 IVLGLGNLTVLYCMKSNLINSVSNITPMNLHVLDVIVICGICPLTIVILLSS-----106
Db 50 IILGVGNL-ALIIILKQKEMRNVTNIIIVNLSFSDLLVAVMCLPFTFYVYTLMDHWVFG 108
QY 107 ---LESNTALICCFHEACVSFASVSTAINFAITL---DRYDISV-----KPNRILTMG 155
Db 109 ETMCKLNPFFVQC-----VSITVSFSLVLAVERHQLIINPRGWRPNRRHAYIG 157
QY 156 RAVMLMISIWIFSFPSFLPIPIEVNPFSLQSGNTWENKTLIC-----VSTNEYYTEL-GM 209
Db 158 ITV-----IWLAVASSL-PFV---IYQIUTDEPFQNVSLAAKDKVYCFDQFPSPDSHRL 208
QY 210 YYHLLVQIPIFFFTVVVYMLITYTKILQALNIRIGTRFSTGQKKKARKKKTISLTQHEAT 269
Db 209 SYTTLLLVLYQFGLPCFIFICYFYI-----IRLKR-----240
QY 270 DMSOSSGRNVVGVRTSVSVIITALRAVKRHRERERQKRVFRMSILLIISTFLLCWTPI 329
Db 241 -----NNMMDKIRDS-----KYRSETKRINIMLLSIVVAFVACWLPL 278
QY 330 SVLNTT-----IILGPSDLLVKRLCLFVLVMAYGTTIFPHLLYAFTRQKFQKVLK---379
Db 279 TIFNTVFDWNHQIATCNHLL--FLCHLTAMISTCV-NPIFYGLNKNFQRLDQFFFN 335
QY 380 -----SKMKRV--VSIVEADPLPNNAVITHNSWIDPFRKNKKTTFDSE 420

```

Db 149 TPRRAALISLWTLGFLISIPPM-----GWRTPEDRSDPDACITSKDHGYIYST 200
Qy 207 LQMYHLLVQPIFFFTVVMILITTKILOALNIRIGTRFTSGOKKARKKKTISLTTOH 266
Db 201 FGAFT-----IPL-----LLMLVYGRIFRAARFRI-----RKTVKKVEKTGADTRH 242
Qy 267 EATDMSQ-----SSGGRNVVFGVRTSVSVIILARRAVK----- 299
Db 243 GASPAQPKKSVNGSGSRNRLGVSAGALCANGAVROGDDGAALAEVIEVHRVGNK 302
Qy 300 -----RHRERRQRKVRMSL-----LIISTFLLCWTP 328
Db 303 EHLPLPSEAGPTPCAPASFERKNERNAEAKR--KMLAREPKVTKTLGIMGTFLCWLP 360
Qy 329 ISVLNTTI-LCLGPSDLLVKLRCLCLVMAVGTTHPHLLYAFTRQKFKVLSKMK 383
Db 361 FFIVALVLPFCSESSCHMPTLLGAIINWLGYNSLLNPVIYAYFNKDFQNAFKKI 416

RESULT 2
US-09-962-646-15
; Sequence 15, Application US/09962646
; Patent No. US20020103123A1
; GENERAL INFORMATION:
; APPLICANT: GERALD, CHRISTOPHE P.G.
; APPLICANT: WEINSHANK, RICHARD L.
; APPLICANT: WALKER, MARY W
; APPLICANT: BRANCHEK, THERESA
; TITLE OF INVENTION: MODIFYING FEEDING BEHAVIOR, COMPOUNDS USEFUL IN SUCH METHODS, AND
; FILE REFERENCE: 1795/46166B2A
; CURRENT APPLICATION NUMBER: US/09/962,646
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/200,673
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 08/566,096
; PRIOR FILING DATE: 1995-12-01
; PRIOR APPLICATION NUMBER: 08/349,025
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-962-646-15

Query Match 9.5%; Score 208; DB 10; Length 384;
Best Local Similarity 21.9%; Pred. No. 1.5e-11;
Matches 78; Conservative 68; Mismatches 108; Indels 102; Gaps 15;

Qy 52 IVLGLSMLTVLVLYCMKSNLINSVNTITNHLVLDVVICGCIPLTVILLIS----- 106
Db 51 IILVSGNL-ALLIILKOKEMRNVTNLIIVNLSFSLDLVAIMCLPLTFVTLMDHWFV 109
Qy 107 ---LESNTALICCFHEACVSASVSTAINFAITL---DRYDISV-----KPANRLTWG 155
Db 110 EAMCLNPFVQC-----VSITISIFSLVIAVERHQIINPGRPNRHHAYV 158
Qy 156 RAVMLMISITWIFSFSLPIETVNFSSLSQGNWTENKTL-----LCVSTNEYTEL-GM 209
Db 159 IAV-----IWLAVASSL-PEL-----IYQVMTDEPFQNTLDAYKDYVCFOFPDSHRL 209
Qy 210 YHLLVQPIFFFTVVMILITTKILOALNIRIGTRFTSGOKKARKKKTISLTQHEAT 269
Db 210 SYTTILLVQVFGPLCFICFYKFIY-----IRLKRNNMMDK----- 247
Qy 270 DMSQSSGGRNVVFGVRTSVSVIILARRAVKRRRRERQRKVRMSLLIISTFLLCWTP 329
Db 248 -----MRDNKRSSTKRNIMLLSIVVAFVAVCWLP 279
Qy 330 SVLNTT-----ILCLGPSDLLVKLRCLCLVMAVGTTHPHLLYAFTRQKFKV 379

Db 280 TIENTVFDNHHQIATCNHNL--FLLCHLTAMISTCV-NPIFYGLNKNFORDLQ 332

RESULT 3
US-10-052-589-2
; Sequence 2, Application US/10052589
; Patent No. US2002013382A1
; GENERAL INFORMATION:
; APPLICANT: Perez, Dianne
; APPLICANT: Zuscik, Michael
; TITLE OF INVENTION: Model systems for neurodegenerative and cardiovascular disorder
; FILE REFERENCE: 26473/04200
; CURRENT APPLICATION NUMBER: US/10/052,589
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/568,255
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Mesocricetus sp.
US-10-052-589-2

Query Match 9.5%; Score 208; DB 12; Length 497;
Best Local Similarity 19.3%; Pred. No. 2e-11;
Matches 85; Conservative 99; Mismatches 163; Indels 94; Gaps 17;

Qy 22 DIDDIN-TNMQPLSYPLSFVSLTGLFMLEIVLG-----LGSNLTVLVLYCKMSNLI 73
Db 19 ELKDANFTGPQNTSSNTLPQLDVTTRATISVGLVLCAGFILFAIVGNILVLSVACNRH--L 76
Qy 74 NSVSNITMNLHLVDVVICGCIPLTVILLISLESNTALICCFHEACVSASVSTAINV 133
Db 77 RPTNYFIVNLAIADLLSFTVLPFSATLEVLYGVWVLRIFCDIWAADVLCCTASILSL 136
Qy 134 FAITLDYDISVKPANR---ILTMGRVAMLMISWIFFSFFSLPIFIEVNFPSLQSGTW 190
Db 137 CAISIDRY-IGRVSLQYPTLVTRKKAIALLSVWLSTVISIGPLL-----GW 184
Qy 191 E---NKTLLCVSTNE---YTELGMYYHLVQIPIFFFTVVMILITTKILOALNIRI 242
Db 185 KEPAPNDDEKCVTEPEFYALFSSLSGSFYIPLAVLMVCRVYIVAKRTKNLEA----- 239
Qy 243 GTRFSTGOKKARKKKTISLTQ-----HEATDMSQSSGGRNVVFGVRTSVSVIILARRAV 298
Db 240 -----GVNKEMSNSKELTRIHSKNFHEDTILSSTKAKGHN-----PRSSIAVKLF----- 284
Qy 299 KHRERRERQRKVRMSLLIISTFLLCWTPISV---LNTTILCLGPSDLLVKLRCLFLVM 355
Db 285 -----KFSREKKAATLGIIVGMFILCWLPFFIALPLGSLFSLTKPPDAVFKV---VFWL 336
Qy 356 AGYTTIFHPLLYAFTRQKFKV-----KSMKKRVSVIVEADPLPNNAVTHN 403
Db 337 GFNSCLNPIIYPCSSKEFKRAFMRILCQCRCSGRRRRRRRLGAC-----AYYR 387
Qy 404 SW-----IDPKRNKKITFEDS 419
Db 388 PWTGRGSLERSQSRKDSLODS 408

RESULT 4
US-09-951-622-10
; Sequence 10, Application US/09951622
; Patent No. US20020106734A1
; GENERAL INFORMATION:
; APPLICANT: Daniel R. Soppet et al.
; TITLE OF INVENTION: ADRENERGIC RECEPTOR
; FILE REFERENCE: PF128D2C1
; CURRENT APPLICATION NUMBER: US/09/951,622
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 09/339,244
; PRIOR FILING DATE: 1999-06-24

GenCore version 5.1.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 16:01:23 ; Search time 12 Seconds
(without alignments)
586.077 Million cell updates/sec

Title: US-09-845-721-2

Perfect score: 2192

Sequence: 1 MCFSPILEINQWSESNITVR.....ITFEDSEIRKRLVQVYTD 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | ID | Description |
|------------|-------|---------------|--------|-------------------------------------|
| 1 | 222 | 10.1 | 451 10 | US-09-993-844-6 Sequence 6, Appli |
| 2 | 208 | 9.5 | 384 10 | US-09-962-646-15 Sequence 15, Appli |
| 3 | 208 | 9.5 | 497 12 | US-10-052-589-2 Sequence 2, Appli |
| 4 | 207.5 | 9.5 | 517 10 | US-09-951-622-10 Sequence 10, Appli |
| 5 | 204 | 9.3 | 384 10 | US-09-771-956-2 Sequence 2, Appli |
| 6 | 203 | 9.3 | 382 10 | US-09-771-956-3 Sequence 3, Appli |
| 7 | 201 | 9.2 | 353 10 | US-09-885-478-4 Sequence 4, Appli |
| 8 | 197.5 | 9.0 | 390 12 | US-09-910-411-2 Sequence 2, Appli |
| 9 | 197.5 | 9.0 | 390 12 | US-10-052-193-2 Sequence 2, Appli |
| 10 | 195 | 8.9 | 390 12 | US-10-005-010-4 Sequence 4, Appli |
| 11 | 195 | 8.9 | 481 8 | US-08-681-219-31 Sequence 31, Appli |
| 12 | 195 | 8.9 | 481 10 | US-09-919-437-72 Sequence 72, Appli |
| 13 | 195 | 8.9 | 518 10 | US-09-804-551B-8 Sequence 8, Appli |
| 14 | 193.5 | 8.8 | 370 12 | US-10-044-592-12 Sequence 12, Appli |
| 15 | 193.5 | 8.8 | 370 12 | US-10-044-592-74 Sequence 74, Appli |
| 16 | 192.5 | 8.8 | 413 10 | US-09-811-286-2 Sequence 2, Appli |
| 17 | 192.5 | 8.8 | 413 10 | US-09-951-622-12 Sequence 12, Appli |
| 18 | 192.5 | 8.8 | 413 10 | US-09-993-844-2 Sequence 2, Appli |
| 19 | 192.5 | 8.8 | 413 10 | US-09-990-596-1 Sequence 1, Appli |

| | | | | |
|----|-------|-----|--------|-------------------------------------|
| 20 | 192.5 | 8.8 | 413 10 | US-09-922-494-13 Sequence 13, Appli |
| 21 | 192.5 | 8.8 | 431 9 | US-10-029-009-8 Sequence 8, Appli |
| 22 | 192.5 | 8.8 | 451 9 | US-10-029-009-20 Sequence 20, Appli |
| 23 | 191 | 8.7 | 370 10 | US-09-993-844-3 Sequence 3, Appli |
| 24 | 191 | 8.7 | 431 10 | US-09-826-508-22 Sequence 22, Appli |
| 25 | 191 | 8.7 | 431 10 | US-09-899-532-2 Sequence 2, Appli |
| 26 | 187.5 | 8.6 | 466 10 | US-09-951-622-11 Sequence 11, Appli |
| 27 | 187 | 8.5 | 382 10 | US-09-993-844-5 Sequence 5, Appli |
| 28 | 186 | 8.5 | 353 10 | US-09-895-686-2 Sequence 2, Appli |
| 29 | 185.5 | 8.5 | 372 10 | US-09-966-871-80 Sequence 80, Appli |
| 30 | 185.5 | 8.5 | 372 12 | US-10-039-645-80 Sequence 80, Appli |
| 31 | 183.5 | 8.4 | 372 10 | US-09-214-904-4 Sequence 4, Appli |
| 32 | 182.5 | 8.3 | 353 10 | US-09-925-776-2 Sequence 2, Appli |
| 33 | 182.5 | 8.3 | 353 10 | US-09-885-478-28 Sequence 28, Appli |
| 34 | 182.5 | 8.3 | 375 10 | US-09-962-646-17 Sequence 17, Appli |
| 35 | 182.5 | 8.3 | 422 10 | US-09-885-478-2 Sequence 2, Appli |
| 36 | 182.5 | 8.3 | 422 10 | US-09-885-478-26 Sequence 26, Appli |
| 37 | 182.5 | 8.3 | 422 10 | US-09-885-478-27 Sequence 27, Appli |
| 38 | 182 | 8.3 | 370 10 | US-09-823-114-21 Sequence 21, Appli |
| 39 | 180 | 8.2 | 460 10 | US-09-989-861-18 Sequence 18, Appli |
| 40 | 179 | 8.2 | 369 10 | US-09-823-114-9 Sequence 9, Appli |
| 41 | 178.5 | 8.1 | 372 10 | US-09-823-114-8 Sequence 8, Appli |
| 42 | 177 | 8.1 | 540 10 | US-09-804-551B-2 Sequence 2, Appli |
| 43 | 176 | 8.0 | 471 10 | US-09-989-861-17 Sequence 17, Appli |
| 44 | 175.5 | 8.0 | 529 10 | US-09-951-622-2 Sequence 2, Appli |
| 45 | 174.5 | 8.0 | 367 10 | US-09-823-114-23 Sequence 23, Appli |

ALIGNMENTS

RESULT 1

US-09-993-844-6

; Sequence 6, Application US/09993844

; Patent No. US00020106739A1

; GENERAL INFORMATION:

; APPLICANT: Oakley, Robert H.

; APPLICANT: Barak, Lawrence S.

; APPLICANT: Laporte, Stephane A.

; APPLICANT: Caron, Marc G.

; TITLE OF INVENTION: Modified G-Protein Coupled Receptors

; FILE REFERENCE: 033072-026

; CURRENT APPLICATION NUMBER: US/09/993,844

; CURRENT FILING DATE: 2001-11-05

; PRIOR APPLICATION NUMBER: US 60/245,772

; PRIOR FILING DATE: 2000-11-03

; PRIOR APPLICATION NUMBER: US 60/260,363

; PRIOR FILING DATE: 2001-01-08

; NUMBER OF SEQ ID NOS: 81

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 451

; TYPE: PPT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: amino acid sequence of 5HT1AR-V2R chimera

; OTHER INFORMATION: expressed from the pEarB-1/5HT1AR vector

US-09-993-844-6

Query Match 10.1%; Score 222; DB 10; Length 451;

Best Local Similarity 21.9%; Pred. No. 9.4e-13;

Matches 91; Conservative 70; Mismatches 153; Indels 102; Gaps 15;

QY 38 LSPQV---SLTGKLMLEIVLGLSNLTVLVLYCMKNSLINSVSNITMNLHVLVDIVCVG 94

Db 33 VSTQVTSLLGLTGLIFCAVLG---NACVAAIALERSLQN-VANYLIGSLAVTDLMSVSL 88

QY 95 CIPITIVILLLESNTALICCFHEACVSPASVSTAINVFATIDRDYDISVKPANRI--L 152

Db 89 VLPMAALYQVLNKNWTLGQVTCDLFIADLVLCCTTSILHLCAIALDRYWAITDPIDYNKR 148

QY 153 TMGRVAVMLTSIWFSPFSLPIPIEFVFNFSLOSQNTWENKT--LLCVSTNEY-----YTE 206

```

Db 118 AVLCHPFTLIPNLMEFIQEVIC---RAAAYFGLSVSVSTFNLVAISIERYSIAICNPL 174
QY 148 ANRI-LTNGRAVLMISIWIPFSFLIPFIEVN---FFSL-----QSGNTWENKTLL 196
Db 175 XSRVQTRSHAYRVIAATWVLSI-IMIPYLVNKTVTFFPMKDRRVGHQCRVLWPSK--- 230
QY 197 CVSTNEYVTELGMYHLLVQIPFFFTVVMLTYTKILOALNIRIGTRFSTGQKKARK 256
Db 231 -----QVOQAWIVLTLTFFIPGVVMIVAYGLISRELYRGIOFEMDLNKEAKHK 281
QY 257 KKTISLTQHEATD---MSQSSGGRNVVFGVRTSVSVIALLRRA-VKRHRERRERQKV 311
Db 282 NGVSTPTTIPSGDEGDCYIQVTKRNTMENSTLTPSVCTKMDRARINNSEAKLMKKRV 341
QY 312 FRMSLLIISTFLLCWTPISVLNT--TILCLGPSDLLVKRLCFL-VMAYGTFIFHPLIYA 368
Db 342 IRMLIVIVAMFFICWPIFVANTWKADELAFNTLTGAPISFIHLSTYTSACVNPLIYC 401
QY 369 FTRQKFKVLKSKMKKRVSVIVEADPLPNNAVIHNSWIDPKENKKTRED 418
Db 402 FMNKRFRKFLGTF-----SSCIRPCNFRDQD 431

```

Search completed: December 4, 2002, 16:03:12
 Job time : 20 secs

```

QY 197 CVSTNEYTELGMYYHLLVOIDPIFFTVVVMALITYTKILQALNIRGTFRPSTGQKKKARK 255
Db 231 -----QVOQAWYVLLLTILFFIPGVVMIVAYGLISRELYRGIQFEMDLNKEAKAHK 281
QY 257 KKTISLTQHEATD---MSQSSGGRNVFVGRTSVSVIIALRRA-VKRHRERERQKRV 311
Db 282 NGVSTPTTIPSGDEGGCYIQVTKRRNTMEMSFLTSPVCTKMDRARNINSEAKLMAKRV 341
QY 312 FRMSLLIISFLLCWTPISVINT--TILCIGPSDLLVKLRLCFL-VMAYGTTIFHPLLYA 368
Db 342 IRMLIVAMFFICWPIFVANTWKAFDELSAFNTLGAPISFIHLLSYTSACVNPILIYC 401
QY 369 FTRQKFKQVLKSKMKKRVSVIADPLPNNAVHNSWIDPKRKNKKTTFED 418
Db 402 FMNKRERKAFLGTF-----SSCIKPCRNFROTDED 431

RESULT 15
US-09-076-510-7
; Sequence 7, Application US/09076510
; Patent No. 6376198
; GENERAL INFORMATION:
; APPLICANT: Kopin, Alan S.
; APPLICANT: Beinborn, Martin
; TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO
; TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSeq version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,510
; FILING DATE: 12 MAY 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/570,157
; FILING DATE: 11-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/109002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/428-0200
; TELEFAX: 617/438-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-076-510-7

Query Match 11.4%; Score 249.5; DB 4; Length 453;
Best Local Similarity 22.7%; Pred. No. 1e-13;
Matches 93; Conservative 84; Mismatches 172; Indels 61; Gaps

QY 32 QPLSYPLSFQVSLTGFLMLEIVLGLSGNLVLYVLYCMKSNLINSVNIITMNLHLVDVII 91
Db 60 KPSPSDLNLMVRVMYVSVFLLSVFGNTLIIIVLVNKR--LRTITNSFLLSLALSDLMV 117
QY 92 CVGCIPLTIVILLLSLSNTALICCCFHEACVSFASVSTAINVF---AITLDRYDISVKP- 147

```



```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSeq version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,510
; FILING DATE: 12 MAY 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/570,157
; FILING DATE: 11-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/109002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/428-0200
; TELEFAX: 617/438-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 443 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-076-510-6

Query Match 11.7%; Score 257; DB 4; Length 443;
Best Local Similarity 23.6%; Pred. No. 2.3e-14;
Matches 102; Conservative 84; Mismatches 131; Indels 112; Gaps 21;

QY 6 ILBNMQSENIIVRDDIDDININMY-----QP-LSYPLSFQVSLTGFLMLEIVLGLSN 59
DB 18 VVDSLMLNGSNITPPCELGLENETLFCLOPQPSKEWQALQILLYSIIFLSVLGNTLV 77
QY 60 LTVLVLYCMKSNLINSVNIITMNLHVLVDVVICVGCIPLTIVILLLESNTALICCFHE 119
DB 78 ITVLI-----RNKRMRTVTNIFLLSLAVSDLMCLFC-MCPFNLPNLLKDFIGSAVC---K 129
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRAVMLMISWIFSF----- 169
DB 130 TTTYFMGTSVSTFNLSLERYGAICRPLQSRVWQTKSHALKVIAATWCLSFMTMP 189
QY 170 ---FSLPIPIEVN-----FSLQSG---NTWENKTLICVSTNEYVELGMYVHLLVQ 216
DB 190 YPIYSNLEVPFTKNNQNTANMCRFLPSDAMQOSWQTFLLI----- 230
QY 217 IPIFFTVVVMITYTKILQALNIRIGTRFSTGQKKARKKK-TISLTTQHEATD----- 270
DB 231 --LFLLPGLVIMVAYGLI--SLEYQIKFDASQKSAKEKRPSTGSTRYEDSDGCLYQ 286
QY 271 -----MSQSGGRNVFVRTSVSV---IALRRVKKRRERERQKRVFMSL 316
DB 287 KSRPPKLELQQLSGSGSGSR-LNRRSSSAANLIA-----KKRVIRMLI 331
QY 317 LIISTELLQWTPISVLT-----TI-----LCUGPSDLLVLRCLFLVMAYGTTIFHPLL 366
DB 332 VIVVFLCWMPLFSANAWRAYDTVSAEKLSTGTPISFI-----LLSYTSSCVNPII 384
QY 367 YAFTRQKFQ 375
DB 385 YCFMKNRFR 393

RESULT 11
US-08-570-157-5
; Sequence 5, Application US/08570157
; Patent No. 5750353
; GENERAL INFORMATION:
; APPLICANT: Kopin, Alan S.
; APPLICANT: Beinborn, Martin
```

```
;
; TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO
; TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,157
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/109001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 428 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-157-5

Query Match 11.5%; Score 253; DB 1; Length 428;
Best Local Similarity 22.8%; Pred. No. 4.8e-14;
Matches 95; Conservative 85; Mismatches 132; Indels 104; Gaps 18;

QY 15 SNITVRDDIDDININMY-----QPLSYPLSFQVSLTGFLMLEIVLGLSN 59
DB 12 SNITPPCELGLENETLFCLOPQPSKEWQPAVQILLYSLIFLSV-----LGNT 60
QY 60 LTVLVLYCMKSNLINSVNIITMNLHVLVDVVICVGCIPLTIVILLLESNTALICCFHE 119
DB 61 LVITVL--INRKMRTVTNIFLLSLAVSDLMCLFCMPFNLPNLLKDFIGSAVC---K 115
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRAVMLMISWIFSF----- 169
DB 116 TTTYFMGTSVSTFNLSLERYGAICRPLQSRVWQTKSHALKVIAATWCLSFMTMP 175
QY 170 ---FSLPIPIEVNFFSLQSCNTWENKTL-LC--VSTNEYVELGMYVHLLVQIPFFFT 223
DB 176 YPIYSNLEVPFTKNN-----NOTANMCRFLPNVDV---MQOSWHTFLLILLFLIP 221
QY 224 VVYMLITYTKILQALNIRIGTRFSTGQKKARKKKTISLTT-QHEATD----- 270
DB 222 GIYVMVAYGLI--SLEYQIKFEASQKSAKEKRPSTSSGKYEDSDGCLYLOKTRPRK 279
QY 271 -----MSQSGGRNVFVRTSVSVIIALRRVKKRRERERQKRVFMSLIIISTFLLC 325
DB 280 LEURLQSTGSSSRANRIRSNSSAANLMA-----KKRVIRMLIVIVLWFLFLC 325
QY 326 WTPISVLTILCLGPSDLLVLRCL-----FLVMAYGTTIFHPLLYATROKFQ 375
DB 326 WMPIFSANA---WRAYDTASAERRLSGTPISFILLSYTSSCVNPIIYCFMKNRFR 378

RESULT 12
US-08-029-170-31
; Sequence 31, Application US/08029170
; Patent No. 6169173
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 444 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-029-170-14

Query Match 12.2%; Score 266.5; DB 4; Length 444;
Best Local Similarity 23.8%; Pred. No. 3.5e-15;
Matches 102; Conservative 85; Mismatches 131; Indels 111; Gaps 20;

QY 6 ILEINMQSESNTVRDDIDDINTNMY-----QP-LSYPLSFQVSLTGFLMLEIVLGLGNS 59
DB 18 VVDSLLMGNSNITPPCELGLENELFCLDQPOPKSEWQSALQILLYSIIFLLSVLGLNTLV 77
QY 60 LTVLVLYCMKSNLNSVNIITMNLHVLVDVICGCIPLTIVILLLSLESNTALICCPHE 119
DB 78 ITVLI-----RNKRMRTVTNIFLLSLAVSDMLCLFCMPFNLPNLLKDFIFGSAVC---K 130
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRVLMISWIFSF----- 169
DB 131 TTTVMGTSVSVSTFNVAISLERYGAICRPLQSRVWTKSHALKVIAATWCLSTMTMP 190
QY 170 ---FSFLPFTEVN-----FFSLQSG---NTWENKTLICVSTNEYYTELGMYHLLVQ 216
DB 191 YPIYSLNVPFTKNNQNTANMCRFLPSDAMQSQWTFLLI----- 231
QY 217 IPIFFFTVVMLIYTKILOALNIRIGTRFSTGOKKARKK-TISLTQHEATD----- 270
DB 232 --LFLPGLVVMVAYGLI--SLEYQGKFDASQKSAKEKPKSTGSTRYEDSDGCYIQ 287
QY 271 -----MSQSSGGRNVFGRVTSV--IIALRRVAKRHRERERQKRVFMSL 316
DB 288 KSRPPRKLQLQSLSGSGSR-LNRRSSSAANLIA-----KKRVIRMLI 332
QY 317 LIISTFLCWTPIISVLNT-----TI-----LCLGPSDLLVKRLCFLVMYGTTFIFHPLL 366
DB 333 VIVVFLCWMPIFSANAWRAYDVSAEKLHSGTPISFI-----LLSYTSSCVNPDI 385
QY 367 YAFTRQRFQ 375
DB 386 YCFMKNRFR 394

RESULT 9
US-08-570-157-6
Sequence 6, Application US/08570157
Patent No. 5750353
GENERAL INFORMATION:
APPLICANT: Koplin, Alan S.
APPLICANT: Beinborn, Martin
TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO
TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,157

FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/109001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-157-6

Query Match 11.7%; Score 257; DB 1; Length 443;
Best Local Similarity 23.8%; Pred. No. 2.3e-14;
Matches 102; Conservative 84; Mismatches 131; Indels 112; Gaps 21;

QY 6 ILEINMQSESNTVRDDIDDINTNMY-----QP-LSYPLSFQVSLTGFLMLEIVLGLGNS 59
DB 18 VVDSLLMGNSNITPPCELGLENELFCLDQPOPKSEWQSALQILLYSIIFLLSVLGLNTLV 77
QY 60 LTVLVLYCMKSNLNSVNIITMNLHVLVDVICGCIPLTIVILLLSLESNTALICCPHE 119
DB 78 ITVLI-----RNKRMRTVTNIFLLSLAVSDMLCLFCMPFNLPNLLKDFIFGSAVC---K 129
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRVLMISWIFSF----- 169
DB 130 TTTVMGTSVSVSTFNVAISLERYGAICRPLQSRVWTKSHALKVIAATWCLSTMTMP 189
QY 170 ---FSFLPFTEVN-----FFSLQSG---NTWENKTLICVSTNEYYTELGMYHLLVQ 216
DB 190 YPIYSLNVPFTKNNQNTANMCRFLPSDAMQSQWTFLLI----- 230
QY 217 IPIFFFTVVMLIYTKILOALNIRIGTRFSTGOKKARKK-TISLTQHEATD----- 270
DB 231 --LFLPGLVVMVAYGLI--SLEYQGKFDASQKSAKEKPKSTGSTRYEDSDGCYIQ 286
QY 271 -----MSQSSGGRNVFGRVTSV--IIALRRVAKRHRERERQKRVFMSL 316
DB 287 KSRPPRKLQLQSLSGSGSR-LNRRSSSAANLIA-----KKRVIRMLI 331
QY 317 LIISTFLCWTPIISVLNT-----TI-----LCLGPSDLLVKRLCFLVMYGTTFIFHPLL 366
DB 332 VIVVFLCWMPIFSANAWRAYDVSAEKLHSGTPISFI-----LLSYTSSCVNPDI 384
QY 367 YAFTRQRFQ 375
DB 385 YCFMKNRFR 393

RESULT 10
US-09-076-510-6
Sequence 6, Application US/09076510
Patent No. 6376198
GENERAL INFORMATION:
APPLICANT: Koplin, Alan S.
APPLICANT: Beinborn, Martin
TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO
TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

QY 170 ---FSELPIEFIEVNFSLQSGNTWENKTLCC--VSTNEYYTELGMYYHLLVQIPIFFFTV 224
Db 196 YPIYSNLVPTNN--NOTGN-----MCRFLPLNDVMQQT--WHFTFLLLILFLIPG 242
QY 225 VVMLTYTKILOALNIRIGTRFSTGOKKARKKTKIS-----LTTQH---- 266
Db 243 IVMWAYGLI--SLEYQIGIKEDAIQKSAKERTKSTGSGPMEDSDGCLYLOKSRHPRKL 300
QY 267 EATDSQSSGGRNVFGVTSVITIALRRVAKRHREREROKRVFRMSLLIISTIFLLCW 326
Db 301 ELRQLSPSSGSRNIRIRSSSTANLMAK-----KRVIRMLIVIVWVFLFLCW 348
QY 327 TPISVLNT-----TI-----LCLGPSDLLVKLRCFLVMAYGTTIFHPLLYAFTROKFQ 375
Db 349 MPIFSANAWRAYDTVSAERHLSGTPIISFI-----LLLSYTSVCVNPITYCFMKNRFR 400

RESULT 7
US-07-937-609-14
; Sequence 14, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/937,609
; FILING DATE: 19920902
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-937-609-14

Query Match 12.2%; Score 266.5; DB 1; Length 444;
Best Local Similarity 23.8%; Pred. No. 3.5e-15;
Matches 102; Conservative 85; Mismatches 131; Indels 111; Gaps 20;
QY 6 ILEINNQSESNITVRDDIDINTNMY-----QP-LSYPLSFQVSLTGFLMLLEIVLGLGSN 59

Db 18 VVDSLMMGNSNITPPCELGENETLFCLOQPOPSKEWQALQILLYSIIIFLLSVLGNLT 77
QY 60 LTFVLVLCMKSNLINSVSNIIITMNLHVLVDIICVGCIPITIVILLLSLESNTALICCFHE 119
Db 78 IIVLI-----RNKRMRTVNIIFLLSLAVSDMLCLFCMPENLPNLLKDFIGSAVC---K 130
QY 120 ACVSFASVSTAINVF---AITLDYDISVKP-ANRI-LTMGRAVMLMIWIPIFSF----- 169
Db 131 TTTTFMGTSVSVSTFNVAISLERYGAICRPLQSRVMQTKSHALKVIAATWCLSFITMP 190
QY 170 ---FSELPIEFIEVNF-----FSLQSG---NWEKNTLLCVSTNEYYTELGMYYHLLVQ 216
Db 191 YPIYSNLVPTNNQNTANMCRFLPLPSDAMQSQWQTFLLI----- 231
QY 217 IPIFFFTVVVMLTYTKILOALNIRIGTRFSTGOKKARKKTK-TISLTTQHEATD----- 270
Db 232 --LFLPLGVIMVAYGLI--SLEYQIGIKEDAIQKSAKERTKSTGSGTRYEDSDGCLYQ 287
QY 271 -----MSQSSGGRNVFGVTSVSV--IIALRRVAKRHREREROKRVFRMSL 316
Db 288 KSRPRKLEQLQSLSGSGSR-LNIRIRSSSSAANLIA-----KRVIRMLI 332
QY 317 LIISTFLLCWTPISVLNT-----TI-----LCLGPSDLLVKLRCFLVMAYGTTIFHPL 366
Db 333 VIVVLEFLCWMPIFSANAWRAYDTVSAERHLSGTPIISFI-----LLLSYTSVCVNP 385
QY 367 YAFTRQKFQ 375
Db 386 YCFMKNRFR 394

RESULT 8
US-08-029-170-14
; Sequence 14, Application US/08029170
; Patent No. 6169173
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,170
; FILING DATE: 19930310
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,609
; FILING DATE: 02-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD


```
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0377 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-835-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544724
; US-08-919-624-3

Query Match      12.9%; Score 282; DB 2; Length 430;
Best Local Similarity 24.1%; Pred. No. 1.6e-16;
Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;

QY 15 SNITVRDDDDIDNNY-----QPLSYPLSFQVSLTGLFMLEIVLGLGNS 59
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 12 SNITSACELGFENETLFCILDRPRPSKEWQPAVQILLYSLIFLSV-----LGNT 60
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 60 LNVLYLCKMNSLINSNIITMNLHVLVDIICVGCIPITIVILLLSLESNTALICCFHE 119
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 61 LVITVL--IRNRMRTVTNIFLLSLAVSDMLCLFCMPFNLPISLLKDFIFGSAVC--K 115
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 120 ACVSPASVSTAINVF---AITLDYDISVKP--ANRI-LTMGRAVLMISWIFSF----- 169
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 116 TTYFMGTSVSVSTFNLAISLERYGAICKPLQSRVWQKSHALKVIAATWCLSTIMTP 175
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 170 ---FSELPIEFIEVNFSLQSGNTWENKTLIC--VSTNEYYTELGMYHLLVQIPIFFTV 224
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 176 YPIYSNLVPFTKNN---NOTGN-----MCRFLLPNDVMQQT---WHTFLLLLILFLIPG 222
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 225 VVMLTYTKILQALNIRIGTRFSTGOKKARKKTTIS-----LTTQH--- 266
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 223 IVWVAYGLI--SLEYQGIKFDQKSAKERKSTGSGGPMEDSDGGLYQKSRHPRKL 280
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 267 EATDMSQSGGRNVFVGVRTSVSIIALRAVYKRREROKRVFRMSLLIISTFLLCW 326
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 281 ELRQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVIVVLFELCW 328
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 327 TPISVLNT-----TI-----LCIGPSDLLVKRLCLFLVMAYGTTIFHPLLYAFTROK 375
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 329 MPIFSANAWRAYDTVSAERHLSGTPISFI-----LLLSYTSNCVNPPIIYCFMKNRFR 380
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 4
US-08-029-170-23
; Sequence 23, Application US/08029170
; Patent No. 6169173
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,170
; FILING DATE: 19930310
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,609
; FILING DATE: 02-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-029-170-23

Query Match      12.9%; Score 282; DB 4; Length 430;
Best Local Similarity 24.1%; Pred. No. 1.6e-16;
Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;

QY 15 SNITVRDDDDIDNNY-----QPLSYPLSFQVSLTGLFMLEIVLGLGNS 59
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 12 SNITSACELGFENETLFCILDRPRPSKEWQPAVQILLYSLIFLSV-----LGNT 60
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 60 LNVLYLCKMNSLINSNIITMNLHVLVDIICVGCIPITIVILLLSLESNTALICCFHE 119
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 61 LVITVL--IRNRMRTVTNIFLLSLAVSDMLCLFCMPFNLPISLLKDFIFGSAVC--K 115
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 120 ACVSPASVSTAINVF---AITLDYDISVKP--ANRI-LTMGRAVLMISWIFSF----- 169
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 116 TTYFMGTSVSVSTFNLAISLERYGAICKPLQSRVWQKSHALKVIAATWCLSTIMTP 175
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 170 ---FSELPIEFIEVNFSLQSGNTWENKTLIC--VSTNEYYTELGMYHLLVQIPIFFTV 224
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 176 YPIYSNLVPFTKNN---NOTGN-----MCRFLLPNDVMQQT---WHTFLLLLILFLIPG 222
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 225 VVMLTYTKILQALNIRIGTRFSTGOKKARKKTTIS-----LTTQH--- 266
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 223 IVWVAYGLI--SLEYQGIKFDQKSAKERKSTGSGGPMEDSDGGLYQKSRHPRKL 280
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 267 EATDMSQSGGRNVFVGVRTSVSIIALRAVYKRREROKRVFRMSLLIISTFLLCW 326
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 281 ELRQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVIVVLFELCW 328
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 327 TPISVLNT-----TI-----LCIGPSDLLVKRLCLFLVMAYGTTIFHPLLYAFTROK 375
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DQ 329 MPIFSANAWRAYDTVSAERHLSGTPISFI-----LLLSYTSNCVNPPIIYCFMKNRFR 380
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 5
US-07-937-609-24
; Sequence 24, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
```

```

Matches 429; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MCFSPLEINQSSNTVRDIDIDINMYQPLSYPLSFQVSLTGFLMLEIVLGLGSLN 60
Db 1 MCFSPLEINQSSNTVRDIDIDINMYQPLSYPLSFQVSLTGFLMLEIVLGLGSLN 60
Qy 61 TVLVYCKMKNLSVNIITMNLHVLVDVLCVGCIPLTIVILLLESNTALICCPHEA 120
Db 61 TVLVYCKMKNLSVNIITMNLHVLVDVLCVGCIPLTIVILLLESNTALICCPHEA 120
Qy 121 CVSFASVSTAINVFAITLDYDISVKPANRILTMGRAVLMISIWIFSFLLPIFEVN 180
Db 121 CVSFASVSTAINVFAITLDYDISVKPANRILTMGRAVLMISIWIFSFLLPIFEVN 180
Qy 181 FFSQSGNTWENKILLCVSTNEYTELGMVYHLLVQIPIPFVTVVMLITVKILQALNI 240
Db 181 FFSQSGNTWENKILLCVSTNEYTELGMVYHLLVQIPIPFVTVVMLITVKILQALNI 240
Qy 241 RIGRFSTGQKKKARKKKTISLTQHEATDMSSQSGGRNVVGVRTSVSVIALLRRRAVKR 300
Db 241 RIGRFSTGQKKKARKKKTISLTQHEATDMSSQSGGRNVVGVRTSVSVIALLRRRAVKR 300
Qy 301 HRERRERQKRVFRMSLLIISTFLLCWTPIISVLTNTILCLGSDLLVLRCLFLVMAYGTT 360
Db 301 HRERRERQKRVFRMSLLIISTFLLCWTPIISVLTNTILCLGSDLLVLRCLFLVMAYGTT 360
Qy 361 IFHPLLAFTRQKQKVLKSKMKRVSVISVEADPLPNNAVIHNSWIDPKRKKKITFEDSE 420
Db 361 IFHPLLAFTRQKQKVLKSKMKRVSVISVEADPLPNNAVIHNSWIDPKRKKKITFEDSE 420
Qy 421 IREKRLVPQVYTD 433
Db 421 IREKRLVPQVYTD 433

```

RESULT 2

```

US-07-937-609-23
; Sequence 23, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/937,609
; FILING DATE: 19920902
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-937-609-23
Query Match 12.98; Score 282; DB 1; Length 430;
Best Local Similarity 24.18; Pred. No. 1.6e-16;
Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;
Qy 15 SNIVRDIIDIDINMY-----QPLSYPLSFQVSLTGFLMLEIVLGLGSLN 59
Db 12 SNITSACELGFENETLFLDRPRSKHQPAVQILLISLILSV-----LGNT 60
Qy 60 LTVLVYCKMKNLSVNIITMNLHVLVDVLCVGCIPLTIVILLLESNTALICCPHE 119
Db 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPFNLPISLLKDFIFGSAVC---K 115
Qy 120 ACVSFASVSTAINVF---AITLDYDISVKP--ANRI-LTMGRAVLMISIWIFSF----- 169
Db 116 TTTVTGTSVSVFNLVAISLERYGATCKPLQSKVMOTKSHALKVIAATWCLSFITMP 175
Qy 170 ---FSFLLPIFEVNFSLQSGNTWENKILLC--VSTNEYTELGMVYHLLVQIPIFFTV 224
Db 176 YPIYSLNVPFTKNN---NOTGN-----MCRFLLPNDVMOOT---WTFLLILFLIPG 222
Qy 225 VMLITVTKILQALNIRIGTRFSTGQKKARKKKTIS-----LTTQH----- 266
Db 223 IVMMVAYGLI--SLEYQGIKFAIQKSAKERKTSTGSSGPMEDSDGCLQKSRHPRKL 280
Qy 267 EATDMSSQSGGRNVVGVRTSVSVIALLRRRAVKRHRERQKRVFRMSLLIISTFLLCW 326
Db 281 ELQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVIVVLFLLCW 328
Qy 327 TPISVLNT-----TI-----LCLGPSDLLVLRCLFLVMAYGTTIFHPLLAFTRQKFO 375
Db 329 MPISANAWRAYDIVSAERHLSGTPISFI-----LLLSYSSCVNPIIIFCNKRFR 380

```

RESULT 3

```

US-08-919-624-3
; Sequence 3, Application US/08919624
; Patent No. 5994097
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,624
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 16:00:33 ; Search time 18 Seconds
(without alignments)
707.785 Million cell updates/sec

Title: US-09-845-721-2

Perfect score: 2192
Sequence: 1 MCFSPLEINMQSSNITVR.....ITFEDSEIRKRLVQPVVD 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 2168 | 98.9 | 433 | 2 | US-08-919-624-1 |
| 2 | 282 | 12.9 | 430 | 1 | US-07-937-609-23 |
| 3 | 282 | 12.9 | 430 | 2 | US-08-919-624-3 |
| 4 | 282 | 12.9 | 430 | 4 | US-08-029-170-23 |
| 5 | 282 | 12.9 | 450 | 1 | US-07-937-609-24 |
| 6 | 282 | 12.9 | 450 | 4 | US-08-029-170-24 |
| 7 | 266.5 | 12.2 | 444 | 1 | US-07-937-609-14 |
| 8 | 266.5 | 12.2 | 444 | 4 | US-08-029-170-14 |
| 9 | 257 | 11.7 | 443 | 1 | US-08-570-157-6 |
| 10 | 257 | 11.7 | 443 | 4 | US-09-076-510-6 |
| 11 | 253 | 11.5 | 428 | 1 | US-08-570-157-5 |
| 12 | 253 | 11.5 | 428 | 4 | US-08-029-170-31 |
| 13 | 253 | 11.5 | 428 | 4 | US-09-076-510-5 |
| 14 | 249.5 | 11.4 | 453 | 1 | US-08-570-157-7 |
| 15 | 249.5 | 11.4 | 453 | 4 | US-09-076-510-7 |
| 16 | 240.5 | 11.0 | 449 | 1 | US-08-570-157-1 |
| 17 | 240.5 | 11.0 | 449 | 4 | US-09-076-510-1 |
| 18 | 239 | 10.9 | 371 | 1 | US-08-415-818-6 |
| 19 | 239 | 10.9 | 371 | 2 | US-08-894-236-6 |
| 20 | 239 | 10.9 | 371 | 2 | US-08-919-624-4 |
| 21 | 239 | 10.9 | 371 | 5 | PCT-US96-01444-6 |
| 22 | 223 | 10.2 | 451 | 1 | US-08-570-157-2 |
| 23 | 223 | 10.2 | 451 | 4 | US-09-076-510-2 |
| 24 | 222.5 | 10.2 | 452 | 1 | US-07-937-609-16 |
| 25 | 222.5 | 10.2 | 452 | 4 | US-08-029-170-16 |
| 26 | 222 | 10.1 | 422 | 1 | US-07-817-920-3 |
| 27 | 222 | 10.1 | 422 | 1 | US-08-370-542-3 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 222 | 10.1 | 422 | 1 | US-08-117-006-3 | Sequence 3, Appl1 |
| 29 | 222 | 10.1 | 422 | 1 | US-08-216-594-3 | Sequence 3, Appl1 |
| 30 | 222 | 10.1 | 422 | 1 | US-08-542-358-3 | Sequence 3, Appl1 |
| 31 | 222 | 10.1 | 422 | 2 | US-08-157-185-13 | Sequence 13, Appl |
| 32 | 222 | 10.1 | 422 | 2 | US-08-281-526B-13 | Sequence 13, Appl |
| 33 | 222 | 10.1 | 422 | 3 | US-09-018-351-3 | Sequence 3, Appl1 |
| 34 | 222 | 10.1 | 422 | 4 | US-09-332-837-13 | Sequence 13, Appl |
| 35 | 222 | 10.1 | 422 | 5 | PCT-US93-00149-3 | Sequence 3, Appl1 |
| 36 | 219.5 | 10.0 | 411 | 3 | US-08-817-869-3 | Sequence 3, Appl1 |
| 37 | 219.5 | 10.0 | 411 | 5 | PCT-US95-14377-3 | Sequence 3, Appl1 |
| 38 | 219 | 10.0 | 447 | 1 | US-07-937-609-29 | Sequence 29, Appl |
| 39 | 219 | 10.0 | 447 | 4 | US-08-029-170-29 | Sequence 29, Appl |
| 40 | 218 | 9.9 | 447 | 1 | US-07-978-892A-6 | Sequence 6, Appl1 |
| 41 | 217.5 | 9.9 | 448 | 4 | US-08-570-157-3 | Sequence 3, Appl1 |
| 42 | 217.5 | 9.9 | 448 | 4 | US-09-076-510-3 | Sequence 3, Appl1 |
| 43 | 217 | 9.9 | 513 | 2 | US-08-406-855A-21 | Sequence 21, Appl |
| 44 | 217 | 9.9 | 513 | 3 | US-09-206-898-21 | Sequence 21, Appl |
| 45 | 214 | 9.8 | 453 | 1 | US-07-937-609-26 | Sequence 26, Appl |

RESULT 1

US-08-919-624-1
; Sequence 1, Application US/08919624
; Patent No. 5994097
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW HUMAN G-PROTEIN COUPLED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,624
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0377 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CARDNOT01
; CLONE: 282414
US-08-919-624-1

Query Match 98.9%; Score 2168; DB 2; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.2e-177;

This Page Blank (uspto)

```

XX 08-AUG-1996..
XX
XX 30-JAN-1996; 96WO-US01444.
XX
XX 03-APR-1995; 95US-0415818.
XX 03-FEB-1995; 95US-0383746.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Cascleri MA, Linemeyer DL, MacNeil DJ, Shiao L;
XX Strader C, Tan CP, Weinberg DH;
XX
XX WPI; 1996-371369/37.
XX N-PSDB; AAT36127.
XX
XX DNA mol. encoding neuro:peptide Y Yx receptor - useful in assays to
XX identify cpds which bind to receptor, useful to treat, e.g.
XX obesity, diabetes, cardiac vasospasm and Parkinson's disease
XX
XX Claim 38; Page 49-50; 65pp; English.
XX
XX Mouse neuropeptide Y Yx (NPY Yx) receptor (AAW02099), a novel
XX subtype of NPY, is a G-protein coupled receptor having 7
XX transmembrane-spanning domains. Its amino acid sequence was
XX deduced from a genomic DNA fragment (AAT36127) obt'd. from a mouse
XX cosmid library. Vectors were constructed to allow expression of
XX the murine NPY Yx receptor in mammalian (COS-7) cells. The
XX recombinant receptor, or transformed host cells, can be used to
XX screen for cpds. that modulate the function of the receptor, or
XX modulate the expression of nucleic acids encoding the receptor.
XX Such cpds. are useful for treating a variety of disease
XX conditions.
XX
XX Sequence 371 AA;
XX
Query Match 10.9%; Score 239; DB 17; Length 371;
Best Local Similarity 23.3%; Pred. No. 1.2e-16;
Matches 90; Conservative 60; Mismatches 109; Indels 128; Gaps 17;
QY 52 IVLGSGNLTVLVLYCMKNSLINSVSIITMNLHLVDLVICVCIPITIVILLLS--LES 109
DB 46 LMGIFGNLSLIIIIIPKREQONVTNLIANLSLDILVCVMCIPTVIYTLMDHWVFG 105
QY 110 NTALICCFHEACVSFASVSTAI-NVFAITLDYDISVKPANRLITMGRAVMLMISWIFS 168
DB 106 NT---MCKLTSYVQSVSVSIFSLVLAIERVQLIVNPRGKPRVAHAYWGILWLIS 162
QY 169 -----FFSFLI---PF-----IEVNFSLQSG--NTWENK-TLLCVSTN----EYYTEL 207
DB 163 LTLSPFLSYHLTNEPFFHNLSDPTDIYTHQVACVEIWPSEKLNQLLSTSLFMLQYFVPL 222
QY 208 GMYHLLVQIPIFFTVVVMLTYYTKILQALNIRIGTRESTGQKKKARKKKTISLTQHE 267
DB 223 GF-----ILICYLXIVLCI-----RKRT----- 240
QY 268 ATDMSSGSGRNVVFGVRTSVSVIIALRAVKRHRERER---OKRVFMSLLIISTELL 324
DB 241 -----RQVDRKENKSNLENKRVNMLISIVVTFGA 272
QY 325 CWTPISVLNT-----TILCLGPSDLLVKRLGFLVMAYGTTIFHPLLYAFTRQKFKVL 378
DB 273 CMLPLNIFNVIFDWHYHEMLMSCHDLV--FVCHLI-AMVSTCINPLFYGLNKNFKQDL 329
QY 379 KSKMKRNVSVIVEADPLPNNVAIHNSW 405
DB 330 M-----MLIHWCW 337
```

CC anxiety or panic, withdrawal response produced by chronic cerebral or
CC abuse of drugs or alcohol, stress, stroke, hypoglycaemia, cerebral palsy,
CC spinal chord and head injury, poisoning by neurotoxins, infertility,
CC adenomas, obesity or diabetes. The use of the mutant PHRs provides for
CC the more sensitive detection of PHR agonists or antagonists.
CC Note: The present sequence does not appear in the specification; it
CC has been made by modifying the wild type human CCK-A receptor
CC sequence which is provided in pages 54-55.

XX Sequence 428 AA;

Query Match 11.5%; Score 253; DB 18; Length 428;
Best Local Similarity 22.7%; Pred. No. 4.5e-18;
Matches 94; Conservative 86; Mismatches 134; Indels 100; Gaps 18;

QY 15 SNITVRDIDDINTNMY-----QPISYPLSFQVSLTGMLEIVLGLSN 59
DB 12 SNITPPCELGLENETLFCLDQPRPSKEQPAVQILLYSLIFLSV-----LGNT 60
QY 60 LTVLVLYCMKNSLINSVNIITMNLHVDVLCVGCIPLTIVILLLSLESNTALICCFHE 119
DB 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPNLPNLKDFIFGSVC---K 115
QY 120 ACVSPASVSTAINVF---AITLDYDISVKP--ANRI-LTMGRVLMISIWIFSF----- 169
DB 116 TTTYPMTGTSVSVSTENLVSAISLQRYGAICKPQSRVWQTKSHALKVIAATWCLSFITWP 175
QY 170 ---FSFLPFTIEVNFSSQSGNTWENKTL-LC--VSTNEYVTELMYHLLVQIPFFFT 223
DB 176 YPIYSNLVPFTKNN-----NOTANMCRFLPNV---MQQSWHTFLLLLFLIP 221
QY 224 VVMLITTKILOALNIRIGTFSTGQKKARKKKTISLTT--QHEATD----- 270
DB 222 GIVMVAAYGLI--SLELYQGIKFEASQKSAKERKPTSSGKYBDSGCVLQKTRPPRK 279
QY 271 ---MSQSGGRNVFVGVTSVSVIIALRRVRREREROKRVFRMSLLIISTFLLCWT 327
DB 280 LELRLQSTGSSSRANRIRNSNSAHVSAK-----KRVIRMLIVIVLFLCWM 327
QY 328 PISVLTNTILCLGPSDLLVKRLC-----FLVMAYGTTIFHPLLYAFTROKFO 375
DB 328 PIFSANA---WRAYDTASAERLSTPISFILLLSYTSVCVNPITTCFPMNKRFR 378

RESULT 14
AAB66630
ID AAB66630 standard; protein; 428 AA.

XX AAB66630;
XX
XX
XX 05-APR-2001 (first entry)
XX Human CCK A receptor protein.
XX Cholecystokinin; CCK receptor; purify.

XX Homo sapiens.

XX US6169173-B1.

XX 02-JAN-2001.

XX 10-MAR-1993; 93US-0029170.

XX 07-FEB-1992; 92US-0831248.

XX 01-APR-1992; 92US-0861769.

XX 11-AUG-1992; 92US-0928033.

XX 02-SEP-1992; 92US-0937609.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Wank SA;

XX

DR WPI; 2001-136725/14.

XX New cholecystokinin (CCK) receptor-encoding DNA molecule, useful for
PT producing and purifying human CCK receptor protein to
PT sequenceable-grade homogeneity -

XX Disclosure; Fig 13; 82pp; English.

XX The present invention relates to a cholecystokinin (CCK) receptor
CC protein. The CCK receptor-encoding DNA molecule is useful for
CC expressing and purifying CCK receptor protein to
CC sequenceable-grade homogeneity. The CCK receptor proteins
CC or fragments are useful for obtaining antibodies that can
CC recognize CCK-expressing cells. The transformed eukaryotic cell
CC lines are useful for studying the receptor in an environment
CC similar to its native environment, e.g. in the context of
CC studying the electrophysiology or binding properties of the receptor.
CC The transformed prokaryotic or insect cell line is useful for
CC expressing CCK receptor to produce large amounts of the receptor for
CC immunological purposes or for studying protein structure, e.g.
CC crystallography.

XX Sequence 428 AA;

Query Match 11.5%; Score 253; DB 22; Length 428;
Best Local Similarity 22.8%; Pred. No. 4.5e-18;
Matches 95; Conservative 85; Mismatches 132; Indels 104; Gaps 18;

QY 15 SNITVRDIDDINTNMY-----QPISYPLSFQVSLTGMLEIVLGLSN 59
DB 12 SNITPPCELGLENETLFCLDQPRPSKEQPAVQILLYSLIFLSV-----LGNT 60
QY 60 LTVLVLYCMKNSLINSVNIITMNLHVDVLCVGCIPLTIVILLLSLESNTALICCFHE 119
DB 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPNLPNLKDFIFGSVC---K 115
QY 120 ACVSPASVSTAINVF---AITLDYDISVKP--ANRI-LTMGRVLMISIWIFSF----- 169
DB 116 TTTYPMTGTSVSVSTENLVSAISLQRYGAICKPQSRVWQTKSHALKVIAATWCLSFITWP 175
QY 170 ---FSFLPFTIEVNFSSQSGNTWENKTL-LC--VSTNEYVTELMYHLLVQIPFFFT 223
DB 176 YPIYSNLVPFTKNN-----NOTANMCRFLPNV---MQQSWHTFLLLLFLIP 221
QY 224 VVMLITTKILOALNIRIGTFSTGQKKARKKKTISLTT--QHEATD----- 270
DB 222 GIVMVAAYGLI--SLELYQGIKFEASQKSAKERKPTSSGKYBDSGCVLQKTRPPRK 279
QY 271 ---MSQSGGRNVFVGVTSVSVIIALRRVRREREROKRVFRMSLLIISTFLLC 325
DB 280 LELRLQSTGSSSRANRIRNSNSAANLMA-----KRVIRMLIVIVLFLC 325
QY 326 WTPISVLTNTILCLGPSDLLVKRLC-----FLVMAYGTTIFHPLLYAFTROKFO 375
DB 326 WMPIFSANA---WRAYDTASAERLSTPISFILLLSYTSVCVNPITTCFPMNKRFR 378

RESULT 15

AAW02099

ID AAW02099 standard; Protein; 371 AA.

XX AAW02099;

XX 25-OCT-1996 (first entry)

XX Mouse neuropeptide Y Yx receptor.

XX Neuropeptide Y Yx receptor; NPY Yx; G-protein coupled receptor;

XX obesity; diabetes; cardiac vasospasm; Parkinson's disease.

XX Mus sp.

XX WO9623809-A1.

```
PD 19-JUN-1997.
XX 11-DEC-1996; 96WO-US19958.
XX 03-SEP-1996; 96US-0718047.
PR 11-DEC-1995; 95US-0570157.
XX (NEWB-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX Beinborn M, Kopin AS;
XX WPI; 1997-332726/30.
XX Assay for peptide hormone receptor ligand using mutant forms of the
PT receptors - where changes in second messenger signalling activity
PT indicate that a compound is an agonist
XX Disclosure; Pages 54-55; 88pp; English.
XX This is the human peptide hormone cholecystokinin (hCCK) receptor A.
CC A mutant form of this receptor can be created by substitution of certain
CC aminoacids (AAW29104). The mutant form of this receptor can be used in a
CC a novel method for determining whether a candidate compound is an
CC agonist or an antagonist of a peptide hormone receptor (PHR). The
CC candidate compound is exposed to the mutant form of the PHR that has the
CC ability to amplify the activity of an agonist as compared to the
CC corresponding wild-type receptor. A change in the second messenger
CC signalling activity of the enhanced receptor can be measured to indicate
CC whether the candidate compound is an agonist or an antagonist. The
CC agonists and antagonists can be used for treating or preventing the
CC disorders involving PHRs. They can be used for treating tumours,
CC gastrointestinal disorders, central nervous system disorders, neuroleptic
CC disorders, depression, schizophrenia, disorders of appetite regulatory
CC systems, anxiety or panic, withdrawal response produced by chronic
CC treatment or abuse of drugs or alcohol, stress, stroke, hypoglycaemia,
CC cerebral palsy, spinal chord or head injury, poisoning by neurotoxins,
CC infertility, adenomas, obesity or diabetes. The use of the mutant PHRs
CC provides for the more sensitive detection of PHR agonists or
XX antagonists.
XX SQ Sequence 428 AA;
Query Match 11.5%; Score 253; DB 18; Length 428;
Best Local Similarity 22.8%; Pred. No. 4.5e-18;
Matches 95; Conservative 85; Mismatches 132; Indels 104; Gaps 18;
QY 15 SNTTVRDIDDDINTNMY-----OPLSVPLSFQVSLGFLMLVLVLGSGN 59
DB 12 SNTTPCEGLENETLFCLDQPPRSKEWQPAVQLLYSLIFLSV-----LGNT 60
QY 60 LTVLVLYCMKSNLINSVNIITMNLHVLVDVLCVGCPIPLTIVILLLESNTALICCFHE 119
DB 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPENLIPNLKDFIFGSAVC--K 115
QY 120 ACVSFASVSTAINVF--AITLDYDISVKP-ANRI-LTMGRAVMLMISWIFSF----- 169
DB 116 TTTYFMGTSVSVTFNLVAISLERYGAICKPLQSRVWQTKSHALKVIAATWCLSFITMP 175
QY 170 ---FSFLPIEYVNFESLQSGNTWENKTL-LC--VSTNEYTYELGMYYHLLVOIPFFFT 223
DB 176 YPIYSNLVPTTKN-----NQTANMCRFLPNQV---MQQSGVTFLLILFLIP 221
QY 224 VVYMLITYTKILQALNIRIGTFSTQKKKARKKTKTISLTT-QHEATD----- 270
DB 222 GIVMMVAYGLI--SLEYQIKFEASQKSAKPKFTTSSGKYEDSDGCGYLQKTRPPRK 279
QY 271 -----MSQSGGRNVFGVTSVVIALLRAVKKRERKRVRMSLLIISTFLLC 325
DB 280 LELRLQSLTGSSSRANRIRSNSSAANLMA-----KKRVIRMLIVLWVLFLLC 325
QY 326 WTPISVNLNTTILCGPSDLLAVKLRLC-----FLVMAYGTTIFHPLLAYPTROKFO 375
DB 326 WMPIFSANA---WRAYDTASAEERLSGTPISFILLYSYSSCVNPIIYCFMNRKR 378
```

```
RESULT 13
AAW29104
ID AAW29104 standard; protein; 428 AA.
XX
XX AAW29104;
XX
DT 12-FEB-1998 (first entry)
XX Enhanced CCK-A/gastrin receptor MH21/35.
XX cholecystokinin; CCK; peptide hormone receptor; PHR; neuropeptide;
XX mutant; messenger signal; agonist; antagonist; human; treatment;
XX tumour; gastrointestinal disorder; central nervous system disorder;
XX neurotoxin; substitution; enhanced receptor; hypoglycaemia, MH21/35.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 138 /label= E1380
XX /note= "wild type Glu is replaced by Gln"
XX Misc-difference 303
XX /label= A303H
XX /note= "wild type Ala is replaced by His"
XX Misc-difference 304
XX /label= N304V
XX /note= "wild type Asn is replaced by Val"
XX Misc-difference 305
XX /label= L305S
XX /note= "wild type Leu is replaced by Ser"
XX Misc-difference 306
XX /label= M306A
XX /note= "wild type Met is replaced by Ala"
XX
XX WO9721731-A1.
XX
XX 19-JUN-1997.
XX
XX 11-DEC-1996; 96WO-US19958.
XX
XX 03-SEP-1996; 96US-0718047.
PR 11-DEC-1995; 95US-0570157.
XX
XX (NEWB-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
XX Beinborn M, Kopin AS;
XX
XX WPI; 1997-332726/30.
XX
XX Assay for peptide hormone receptor ligand using mutant forms of the
XX receptors - where changes in second messenger signalling activity
XX indicate that a compound is an agonist
XX
XX Claim 34; Pages -; 88pp; English.
XX
XX This mutant CCK-A receptor M21/35 is derived from the human peptide
XX hormone cholecystokinin (hCCK) receptor A. This enhanced receptor MH21/35
XX is created by the substitution of wild type Glu to Gln at amino acid
XX position 138 and wild type Ala, Asn, Leu, Met to His, Val, Ser, Ala
XX respectively at amino acid positions 303 to 306. This mutant receptor can
XX be used in a novel method for determining whether a candidate compound is
XX an agonist or an antagonist of a peptide hormone receptor (PHR). The
XX candidate compound is exposed to the mutant form of the PHR that has the
XX ability to amplify the activity of an agonist as compared to the
XX corresponding wild-type receptor. A change in the second messenger
XX signalling activity of the enhanced receptor can be measured to indicate
XX whether the candidate compound is an agonist or an antagonist. The
XX agonists and antagonists can be used for treating or preventing disorders
XX involving PHRs. They can be used for treating tumours, gastrointestinal
XX disorders, central nervous system disorders, neuroleptic disorders,
XX depression, schizophrenia, disorders of appetite regulatory systems,
```



```

XX CC The present invention relates to a cholecystokinin (CKK) receptor
XX CC protein. The CKK receptor-encoding DNA molecule is useful for
XX CC expressing and purifying CKK receptor protein to
XX CC sequenceable-grade homogeneity. The CKK receptor proteins
XX CC or fragments are useful for obtaining antibodies that can
XX CC recognize CKK-expressing cells. The transformed eukaryotic cell
XX CC lines are useful for studying the receptor in an environment
XX CC similar to its native environment, e.g. in the context of
XX CC studying the electrophysiology or binding properties of the receptor.
XX CC The transformed prokaryotic or insect cell line is useful for
XX CC expressing CKK receptor to produce large amounts of the receptor for
XX CC immunological purposes or for studying protein structure, e.g.
XX CC crystallography.
XX CC
XX CC Sequence 430 AA;
XX CC
XX CC Query Match 12.9%; Score 282; DB 22; Length 430;
XX CC Best Local Similarity 24.1%; Pred. No. 3.7e-21;
XX CC Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;
XX CC
XX CC QY 15 SNITVRDDIDDINTNMY-----QPLSYPLSFQVSLTGFLMLEIVLGLSN 59
XX CC Db 12 SNITSACELGFENETLFCDDRPRPSKEQPAVQILLYSLIFLLSV-----LGMT 60
XX CC
XX CC QY 60 LTVLVLYCKMSNLINSVNIITMNLHVDVIVCGICPLTIVILLLESNTALICCFHE 119
XX CC Db 61 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPFNLPISLKKDFIGSAVC---K 115
XX CC
XX CC QY 120 ACVSFASYSTAINVF---AITLDYDISVKP--ANRI-LTMGRAVLMISWIFSF----- 169
XX CC Db 116 TTTYEMGTSVSVSTNLVAISLERYGAICKPLQSRVWQTKSHALKAVIAATWCLSTIWP 175
XX CC
XX CC QY 170 ---FSLPIPFIEVNFPSLQSGNTWENKLLC--VSTNEYYTELGMYYHLLVQIPFFFTV 224
XX CC Db 176 YPIYSLNLPVFTKNN---NOTGN-----MCRFLPNVDMQQT---WHTFLLILFLIPG 222
XX CC
XX CC QY 225 VVMLITYTKILOALNIRIGTRFSTGQKKARKKKTIS-----LTTQH--- 266
XX CC Db 223 IVMVAYGLI--SLEYQGIKFDATOKKSAKERKSTSGSGPMEDSDGCYLOKSRHPRKL 280
XX CC
XX CC QY 267 EATDMSQSSGGRNVVFGVRTSVSVIIALRRRAVKRHREREROKRVFRMSLLIISTFLICW 326
XX CC Db 281 ELRQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVIVLFLICW 328
XX CC
XX CC QY 327 TPISVLNT-----TI-----LCLGPSDLLVLRCLFLVMAYGTTTFHPLLYAFTQKQ 375
XX CC Db 329 MPIFSANAWRAYDVTVAERHLSGTPISEI-----LLLSYTSSCVNPITYCFMKNRFR 380
XX CC
XX CC RESULT 8
XX CC AAB66626
XX CC ID AAB66626 standard; protein; 450 AA.
XX CC
XX CC AC AAB66626;
XX CC
XX CC DT 05-APR-2001 (first entry)
XX CC
XX CC DE Guinea pig CCKA receptor protein #2.
XX CC
XX CC KW Cholecystokinin; CKK receptor; purify.
XX CC
XX CC OS Cavia sp.
XX CC
XX CC PN US6169173-B1.
XX CC
XX CC PD 02-JAN-2001.
XX CC
XX CC PF 10-MAR-1993; 93US-0029170.
XX CC
XX CC PR 07-FEB-1992; 92US-0831248.
XX CC
XX CC PR 01-APR-1992; 92US-0861769.
XX CC
XX CC PR 11-AUG-1992; 92US-0928033.

```

```

PR 02-SEP-1992; 92US-0937609.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Wank SA;
XX
XX WPI; 2001-136725/14.
XX
XX New cholecystokinin (CKK) receptor-encoding DNA molecule, useful for
XX PT producing and purifying human CKK receptor protein to
XX PT sequenceable-grade homogeneity -
XX
XX Disclosure; Fig 7; 82pp; English.
XX
XX The present invention relates to a cholecystokinin (CKK) receptor
XX CC protein. The CKK receptor-encoding DNA molecule is useful for
XX CC expressing and purifying CKK receptor protein to
XX CC sequenceable-grade homogeneity. The CKK receptor proteins
XX CC or fragments are useful for obtaining antibodies that can
XX CC recognize CKK-expressing cells. The transformed eukaryotic cell
XX CC lines are useful for studying the receptor in an environment
XX CC similar to its native environment, e.g. in the context of
XX CC studying the electrophysiology or binding properties of the receptor.
XX CC The transformed prokaryotic or insect cell line is useful for
XX CC expressing CKK receptor to produce large amounts of the receptor for
XX CC immunological purposes or for studying protein structure, e.g.
XX CC crystallography.
XX CC
XX CC Sequence 450 AA;
XX CC
XX CC Query Match 12.9%; Score 282; DB 22; Length 450;
XX CC Best Local Similarity 24.1%; Pred. No. 3.9e-21;
XX CC Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;
XX CC
XX CC QY 15 SNITVRDDIDDINTNMY-----QPLSYPLSFQVSLTGFLMLEIVLGLSN 59
XX CC Db 32 SNITSACELGFENETLFCDDRPRPSKEQPAVQILLYSLIFLLSV-----LGMT 80
XX CC
XX CC QY 60 LTVLVLYCKMSNLINSVNIITMNLHVDVIVCGICPLTIVILLLESNTALICCFHE 119
XX CC Db 81 LVITVL--IRNKRMTVTNIFLLSLAVSDMLCLFCMPFNLPISLKKDFIGSAVC---K 135
XX CC
XX CC QY 120 ACVSFASYSTAINVF---AITLDYDISVKP--ANRI-LTMGRAVLMISWIFSF----- 169
XX CC Db 136 TTTYEMGTSVSVSTNLVAISLERYGAICKPLQSRVWQTKSHALKAVIAATWCLSTIWP 195
XX CC
XX CC QY 170 ---FSLPIPFIEVNFPSLQSGNTWENKLLC--VSTNEYYTELGMYYHLLVQIPFFFTV 224
XX CC Db 196 YPIYSLNLPVFTKNN---NOTGN-----MCRFLPNVDMQQT---WHTFLLILFLIPG 242
XX CC
XX CC QY 225 VVMLITYTKILOALNIRIGTRFSTGQKKARKKKTIS-----LTTQH--- 266
XX CC Db 243 IVMVAYGLI--SLEYQGIKFDATOKKSAKERKSTSGSGPMEDSDGCYLOKSRHPRKL 300
XX CC
XX CC QY 267 EATDMSQSSGGRNVVFGVRTSVSVIIALRRRAVKRHREREROKRVFRMSLLIISTFLICW 326
XX CC Db 301 ELRQLSPSSSGSNRINRIRSSSTANLMAK-----KRVIRMLIVIVLFLICW 348
XX CC
XX CC QY 327 TPISVLNT-----TI-----LCLGPSDLLVLRCLFLVMAYGTTTFHPLLYAFTQKQ 375
XX CC Db 349 MPIFSANAWRAYDVTVAERHLSGTPISEI-----LLLSYTSSCVNPITYCFMKNRFR 400
XX CC
XX CC RESULT 9
XX CC AAR38890
XX CC ID AAR38890 standard; protein; 444 AA.
XX CC
XX CC AC AAR38890;
XX CC
XX CC DT 07-FEB-1994 (first entry)
XX CC
XX CC DE Sequence encoded by the rat pancreatic cholecystokinin (CKK) A
XX CC receptor cDNA clone.

```

```

Db 361 IFHPLLAFTRQKQVLKSKMKRRVSVIVEADPLPNNAVIHNSWIDPKRKKITFDSE 420
QY 421 IREKRLVPQVVD 433
    |||| |||||
Db 421 IREKCLVPQVVD 433

RESULT 6
AAR40772
XX AAR40772 standard; Protein; 430 AA.
AC AAR40772;
XX
DT 07-FEB-1994 (first entry)
XX
DE Sequence encoded by the cholecystokinin (CCK) A receptor cDNA in guinea
DE pig gallbladder and pancreas.
XX
KW Cholecystokinin receptor protein; CCK; gastrointestinal receptor.
XX
OS Cavia porcellus.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= glycosylation
FT /note= "see also AAs 12,24,190"
FT 44..67
FT Domain
FT /label= transmembrane domain I
FT 80..103
FT Domain
FT /label= II
FT 118..137
FT Domain
FT /label= III
FT 158..179
FT Domain
FT /label= IV
FT 211..234
FT Domain
FT /label= V
FT 316..336
FT Domain
FT /label= VI
FT 335..374
FT Domain
FT /label= VII
FT Modified-site 249
FT /label= phosphorylation
FT /note= "see also AAs 256,274,292,300,414,416,419"
FT 411
FT Modified-site
FT /label= phosphorylation
XX
XX WO9316182-A.
XX
XX 19-AUG-1993.
XX
XX 28-JAN-1993; 93WO-US00466.
XX
XX 07-FEB-1992; 92US-0831248.
XX 01-APR-1992; 92US-0861769.
XX 11-AUG-1992; 92US-0928033.
XX 02-SEP-1992; 92US-0937609.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Wank SA;
XX
XX WPI; 1993-272886/34.
XX N-PSDB; AAQ47669.
XX
XX Isolated DNA molecule encoding cholecystokinin receptor protein -
XX are purified to isolate cholecystokinin receptor clones and
XX produce anti-cholecystokinin receptor antibodies
XX
XX Example; Figure 6; 110pp; English.
XX
XX The rat pancreatic CCK A receptor cDNA clone encodes a protein with
XX 7 transmembrane domains, and homology with other G-protein receptor
XX superfamily members. There are 4 potential sites of N-linked

```

```

CC glycosylation and sites for serine and threonine phosphorylation.
XX
SQ Sequence 430 AA;
Query Match 12.9%; Score 282; DB 14; Length 430;
Best Local Similarity 24.1%; Pred.No. 3.7e-21;
Matches 101; Conservative 83; Mismatches 127; Indels 108; Gaps 19;
QY 15 SNITVRDDIDDINTNMV-----OPLSYPLSFQVSLTGFLMLEIVLGLGNS 59
    |||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 12 SNITSACELGFENETLFCDDRPRSPKEWQPAVOILLYSLIFLSV-----LGNT 60
    |||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 LTVLVLYCKMSNLINSVNIITMNLHVLVDVVICGCIPLTIVILLLSLESNTALICCFHE 119
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LVITVL--IRNKMRTVTNIFLLSLAVSDMLCLFCMPFNLPISLLKDFIGSAVC---K 115
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 ACVSFASVSAINVF---AITLDYDISVKP-ANRI-LTMGRAVMLMISWISF----- 169
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 TTTYFGTSSVSTFNLVAISLERYGAICKPLOSVMQTKSHALKVIAATWCLSFMTMP 175
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 ---FSFLIPFIEVNFESLQSGNTWENKTLG--VSTNEYVYTELGMVYHLLVOIPFFETV 224
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 YPIYSNLVPFTKNN---NQIGN-----MCRFLLPNDVMQQT---WHFLLLLILFLPG 222
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 VMLITYTKILOALNIRIGTFSTGQKKARKKTKTIS-----LTTQH----- 266
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 IVMWAYGLI--SLEYQGIKFDIAIQKSAKERKTSTGSSGPMEDSDGCVLQKSRHPRKL 280
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 EATDMSOSSGGRNVFGRVTSVVIITALRAVAKRHRERERQKRVFMSLLIISTFLLCW 326
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 ELRQLSPSSSGSNRINRIRSSSSSTANLMAK-----KRVIRMLIVIVLFFLCW 328
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 TPISVLNT-----TI-----LCIGPSDLIVKRLCFLVMAYGTTIFHPLLYAFTROKFO 375
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 329 MPIFSANAWRAYDTVSAERHLSGTPISFI-----LLLSYTSVCVNPPIIYCFMNRFR 380
    | : || : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
AAB66625
ID AAB66625 standard; protein; 430 AA.
XX
AC AAB66625;
XX
DT 05-APR-2001 (first entry)
XX
DE Guinea pig CCKA receptor protein.
XX
KW Cholecystokinin; CCK receptor; purify.
XX
OS Cavia sp.
XX
XX US6169173-B1.
XX
XX 02-JAN-2001.
XX
XX 10-MAR-1993; 93US-0029170.
XX
XX 07-FEB-1992; 92US-0831248.
XX 01-APR-1992; 92US-0861769.
XX 11-AUG-1992; 92US-0928033.
XX 02-SEP-1992; 92US-0937609.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Wank SA;
XX
XX WPI; 2001-136725/14.
XX
XX New cholecystokinin (CCK) receptor-encoding DNA molecule, useful for
XX producing and purifying human CCK receptor protein to
XX sequenceable-grade homogeneity -
XX Disclosure; Fig 6; 82pp; English.
PS

```


agents -
 Example 2; Page 242-244; 34lpp; English.
 The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743 and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X, is substituted for an endogenous residue in IC3 at a position 16 amino acids N-terminal of an endogenous proline in TM6 to form a sequence X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or Ala, and is preferably Lys. When the endogenous residue at this position is Lys, this residue is replaced by His, Arg or preferably Ala. The 15 amino acid stretch between the substituted amino acid and the Pro may be endogenous, non-endogenous, or a mixture of endogenous and non-endogenous residues. The constitutively active GPCRs are useful for identifying antagonists, agonists and partial agonists for use as pharmaceutical agents. The mutant proteins are also useful in research settings for elucidating the roles of the receptors in normal and diseased conditions. Antagonists for a particular GPCR are useful for treating diseases and disorders associated with that receptor. Because the novel mutant GPCRs are constitutively active, they can be used directly for screening of compounds without the need for endogenous ligands. Sequences AAY90643- AAY90677 and AAY90683-Y90687 the mutant human GPCRs of the invention.

Sequence 433 AA;

Query Match 99.6%; Score 2183; DB 21; Length 433;
 Best Local Similarity 99.8%; Pred. No. 5.6e-224;
 Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCFSPLEINMQSESNTVRDDIDDINTNMYQPLSYPLSFQVSLTGFMLLEIVLGLGSLN 60
 Db 1 MCFSPLEINMQSESNTVRDDIDDINTNMYQPLSYPLSFQVSLTGFMLLEIVLGLGSLN 60
 Qy 61 TVLVLYCMKSNLINSVNIITMNLHVLVDVIVCGICPIITVILLLESNTALICCFHEA 120
 Db 61 TVLVLYCMKSNLINSVNIITMNLHVLVDVIVCGICPIITVILLLESNTALICCFHEA 120
 Qy 121 CVSFASVSTAINFAITLDRYDISVKPANRLTMGRAVLMISWIFSFSLPIPFIEVN 180
 Db 121 CVSFASVSTAINFAITLDRYDISVKPANRLTMGRAVLMISWIFSFSLPIPFIEVN 180
 Qy 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVMVLTITTKILQALNI 240
 Db 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVMVLTITTKILQALNI 240
 Qy 241 RIGTRFSTGQKKARKKKTISLTQHEATMSQSGGRNVFVGRTSVSVIALLRAVKR 300
 Db 241 RIGTRFSTGQKKARKKKTISLTQHEATMSQSGGRNVFVGRTSVSVIALLRAVKR 300
 Qy 301 HREREROKRVFRMSLLIISTFLCWPISVLTNTILCLGPSDLLVKLRCLFLVMAYGTT 360
 Db 301 HREREROKRVFRMSLLIISTFLCWPISVLTNTILCLGPSDLLVKLRCLFLVMAYGTT 360
 Qy 361 IFHPLLYAFTROKFKVLKSKMKRVYSIVEADPLPNNAVIHNSWIDPKRKNKKTTFEDE 420
 Db 361 IFHPLLYAFTROKFKVLKSKMKRVYSIVEADPLPNNAVIHNSWIDPKRKNKKTTFEDE 420
 Qy 421 IREKRLVPQVVD 433
 Db 421 IREKRLVPQVVD 433

RESULT 5
 AAY03770
 ID AAY03770 standard; Protein; 433 AA.
 XX
 AC AAY03770;
 XX

DT 11-JUN-1999 (first entry)
 XX Human G-protein coupled receptor (GRech).
 DE
 XX G-protein coupled receptor; Grech; endocrine disorder; neurological;
 KW cardiovascular; human.
 KW
 XX Homo sapiens.
 OS
 XX WO9910491-A1.
 PN
 XX 04-MAR-1999.
 PD
 XX 26-AUG-1998; 98WO-US17690.
 FF
 XX 28-AUG-1997; 97US-0919624.
 PR
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Corley NC, Guegler KJ, Lal P, Shah P;
 PI WPI; 1999-204664/17.
 XX N-PSDB; AAX29296.
 DR
 XX New human G-coupled Receptor (GRech) polypeptide and polynucleotide
 PT - useful as diagnostic reagents and for treatment of endocrine,
 PT cardiovascular and neurological disorders
 PS
 XX Claim 1; Fig 1A-F; 76pp; English.
 XX
 CC This represents a human G-protein coupled receptor (GRech) polypeptide.
 CC Host cells containing a vector comprising the Grech nucleic acid can be
 CC used for the recombinant expression of the protein. Grech polypeptide is
 CC useful in a pharmaceutical composition for treating endocrine,
 CC neurological and cardiovascular disorders. The Grech polynucleotide
 CC complement is useful for detecting a polynucleotide which encodes a Grech
 CC in a sample by hybridising to PCR amplified material, and detecting the
 CC hybridisation complex. Grech antibodies are useful for diagnosis of
 CC conditions/diseases associated with Grech expression, or for monitoring of
 CC patients treated with Grech, agonists, antagonists or inhibitors.
 XX
 SQ Sequence 433 AA;
 Query Match 98.9%; Score 2168; DB 20; Length 433;
 Best Local Similarity 99.1%; Pred. No. 2.2e-222;
 Matches 429; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 MCFSPLEINMQSESNTVRDDIDDINTNMYQPLSYPLSFQVSLTGFMLLEIVLGLGSLN 60
 Db 1 MCFSPLEINMQSESNTVRDDIDDINTNMYQPLSYPLSFQVSLTGFMLLEIVLGLGSLN 60
 Qy 61 TVLVLYCMKSNLINSVNIITMNLHVLVDVIVCGICPIITVILLLESNTALICCFHEA 120
 Db 61 TVLVLYCMKSNLINSVNIITMNLHVLVDVIVCGICPIITVILLLESNTALICCFHEA 120
 Qy 121 CVSFASVSTAINFAITLDRYDISVKPANRLTMGRAVLMISWIFSFSLPIPFIEVN 180
 Db 121 CVSFASVSTAINFAITLDRYDISVKPANRLTMGRAVLMISWIFSFSLPIPFIEVN 180
 Qy 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVMVLTITTKILQALNI 240
 Db 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVMVLTITTKILQALNI 240
 Qy 241 RIGTRFSTGQKKARKKKTISLTQHEATMSQSGGRNVFVGRTSVSVIALLRAVKR 300
 Db 241 RIGTRFSTGQKKARKKKTISLTQHEATMSQSGGRNVFVGRTSVSVIALLRAVKR 300
 Qy 301 HREREROKRVFRMSLLIISTFLCWPISVLTNTILCLGPSDLLVKLRCLFLVMAYGTT 360
 Db 301 HREREROKRVFRMSLLIISTFLCWPISVLTNTILCLGPSDLLVKLRCLFLVMAYGTT 360
 Qy 361 IFHPLLYAFTROKFKVLKSKMKRVYSIVEADPLPNNAVIHNSWIDPKRKNKKTTFEDE 420
 Db 361 IFHPLLYAFTROKFKVLKSKMKRVYSIVEADPLPNNAVIHNSWIDPKRKNKKTTFEDE 420

Db 121 CYSFASVSTAINVFAITLDYDISVKPANRIITMGRVLMISIWTFSPFLIPPIEVN 180
QY 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVVMLITYTKILQALNI 240
Db 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVVMLITYTKILQALNI 240
QY 241 RIGTRFSTGOKKARKKKTISLTTOHEATDMSQSGRNVFVGVRTSVSVIIALRAVR 300
Db 241 RIGTRFSTGOKKARKKKTISLTTOHEATDMSQSGRNVFVGVRTSVSVIIALRAVR 300
QY 301 HRERRRQKRVFRMSLLIISTFLLCWTPISVLTNTTILCLGPSDLLVKKLRCFLVMA YGTT 360
Db 301 HRERRRQKRVFRMSLLIISTFLLCWTPISVLTNTTILCLGPSDLLVKKLRCFLVMA YGTT 360
QY 361 IFHPLLYAFTRQKFQKVLKSKMKRVVSIVEADPLPNNNAVIHNSWIDPKRNKKITFEDSE 420
Db 361 IFHPLLYAFTRQKFQKVLKSKMKRVVSIVEADPLPNNNAVIHNSWIDPKRNKKITFEDSE 420
QY 421 IREKRLVPQVWTD 433
Db 421 IREKRLVPQVWTD 433

RESULT 3
AAM47903
ID AAM47903 standard; Protein; 433 AA.
XX AAM47903;
XX
DT 27-FEB-2002 (first entry)
XX Human GPR22.
DE Human GPR22.
KW Human; GPR22; anorectic; appetite control agent; GPR;
KW G protein-coupled receptor; orphan receptor; antisense gene therapy;
KW mouse; GPR56.
XX
OS Homo sapiens.
XX
PN WO200183550-A2.
XX
PD 08-NOV-2001.
XX
PF 30-APR-2001; 2001WO-GB01874.
XX
PR 03-MAY-2000; 2000US-201418P.
XX
PA (ASTR) ASTRAZENECA AB.
PA (ASTR) ASTRAZENECA UK LTD.
XX
PI Brennand JC, Hart KA;
XX
DR WPI; 2002-066519/09.
DR N-PSDB; ABA05328.
XX
PT Identifying appetite control agent for controlling obesity, comprises
PT screening agonists or antagonists of G protein-coupled receptor, GPR22,
PT and using them as test compounds in appetite control test procedures -
XX
PS Disclosure; Page 15; 21pp; English.
XX
CC The invention relates to identifying an anorectic appetite control agent,
CC comprising screening for agonists and/or antagonists of G
CC protein-coupled receptor GPR22 (an orphan receptor), using one or more
CC agonists and/or antagonists so identified as test compounds in one or
CC more appetite control procedures and selecting an active compound for use
CC as an appetite control agent. An antisense oligonucleotide to the GPR22
CC gene is also useful for controlling obesity using antisense gene therapy.
CC Note: The GPR22 encoding cDNA sequence (ABA05328) and encoded protein
CC (AAM47903) are disclosed, however two DNA sequences described as human
CC (ABA05329) and mouse (ABA05328) GPR56 are given in the sequence listing
CC but are not otherwise referred to in the specification.
XX

SQ Sequence 433 AA;
Query Match 100.0%; Score 2192; DB 23; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.1e-225;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCFSPILEINMSESNTIVRDDIDDINTNMYQPLSPYLSFQVSLTGFLMLEIVLGLGSNL 60
Db 1 MCFSPILEINMSESNTIVRDDIDDINTNMYQPLSPYLSFQVSLTGFLMLEIVLGLGSNL 60
QY 61 TVLVLYCMKSNLINSVSNITMNLHVLVDVICVGCPIPLTIVILLLESNTALICFHEA 120
Db 61 TVLVLYCMKSNLINSVSNITMNLHVLVDVICVGCPIPLTIVILLLESNTALICFHEA 120
QY 121 CYSFASVSTAINVFAITLDYDISVKPANRIITMGRVLMISIWTFSPFLIPPIEVN 180
Db 121 CYSFASVSTAINVFAITLDYDISVKPANRIITMGRVLMISIWTFSPFLIPPIEVN 180
QY 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVVMLITYTKILQALNI 240
Db 181 FFSLSQSGNTWENKTLICVSTNEYTELGMYYHLLVQIPIFFFTVVVMLITYTKILQALNI 240
QY 241 RIGTRFSTGOKKARKKKTISLTTOHEATDMSQSGRNVFVGVRTSVSVIIALRAVR 300
Db 241 RIGTRFSTGOKKARKKKTISLTTOHEATDMSQSGRNVFVGVRTSVSVIIALRAVR 300
QY 301 HRERRRQKRVFRMSLLIISTFLLCWTPISVLTNTTILCLGPSDLLVKKLRCFLVMA YGTT 360
Db 301 HRERRRQKRVFRMSLLIISTFLLCWTPISVLTNTTILCLGPSDLLVKKLRCFLVMA YGTT 360
QY 361 IFHPLLYAFTRQKFQKVLKSKMKRVVSIVEADPLPNNNAVIHNSWIDPKRNKKITFEDSE 420
Db 361 IFHPLLYAFTRQKFQKVLKSKMKRVVSIVEADPLPNNNAVIHNSWIDPKRNKKITFEDSE 420
QY 421 IREKRLVPQVWTD 433
Db 421 IREKRLVPQVWTD 433

RESULT 4
AAY90656
ID AAY90656 standard; Protein; 433 AA.
XX AC AAY90656;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human mutant G protein-coupled receptor GPR22 (F312K).
KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening;
KW agonist; antagonist; mutant; mutein.
OS Homo sapiens.
OS Synthetic.
PN WO200022129-A1.
XX
PD 20-APR-2000.
XX
PF 12-OCT-1999; 99WO-US23938.
XX
PR 13-OCT-1998; 98US-0170496.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Behan DP, Chalmers DT, Liaw CW;
XX
DR WPI; 2000-329165/28.
DR N-PSDB; AAA30722.
XX
PT Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical

PT agents -
PS Example 1; Page 137-139; 341pp; English.
XX
XX
CC The invention relates to constitutively active, non-endogenous versions
CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
CC and AAA30775-A30779). The mutant proteins of the invention contain a
CC mutation in a portion of the protein comprising intracellular loop 3
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC is substituted for an endogenous residue in IC3 at a position 16 amino
CC acids N-terminal of an endogenous proline in TM6 to form a sequence
CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
CC or Ala, and is preferably Lys. When the endogenous residue at this
CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
CC The 15 amino acid stretch between the substituted amino acid and the Pro
CC may be endogenous, non-endogenous, or a mixture of endogenous and
CC non-endogenous residues. The constitutively active GPCRs are useful for
CC identifying antagonists, agonists and partial agonists for use as
CC pharmaceutical agents. The mutant proteins are also useful in research
CC settings for elucidating the roles of the receptors in normal and
CC diseased conditions. Antagonists for a particular GPCR are useful for
CC treating diseases and disorders associated with that receptor. Because
CC the novel mutant GPCRs are constitutively active, they can be used
CC directly for screening of compounds without the need for endogenous
CC ligands. The present sequence represents a human wild-type GPCR referred
CC to in an exemplification of the invention.
XX
SQ Sequence 433 AA;

Query Match 100.0%; Score 2192; DB 21; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.1e-225;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCFSPILEINMQSESNTVRDDIDDINTNMYPQLSPYPLSFQVSLTGFLMLIEIVLGLGSLN 60
DB 1 MCFSPILEINMQSESNTVRDDIDDINTNMYPQLSPYPLSFQVSLTGFLMLIEIVLGLGSLN 60
QY 61 TVLVLYCMKSNLINSVNIITMNLHLVDVLCVGCIPLTIVILLLESNTALICCFHEA 120
DB 61 TVLVLYCMKSNLINSVNIITMNLHLVDVLCVGCIPLTIVILLLESNTALICCFHEA 120
QY 121 CVSPASVSTAINVFAITLDRYDISVKPANRLTMTGRAVLMISWIFSFLIPFTEVN 180
DB 121 CVSPASVSTAINVFAITLDRYDISVKPANRLTMTGRAVLMISWIFSFLIPFTEVN 180
QY 181 FFSLSQGNWENKTLCCVSTNEYTELGMYYHLVQVPIPFVTVVMLITVTKILQALNI 240
DB 181 FFSLSQGNWENKTLCCVSTNEYTELGMYYHLVQVPIPFVTVVMLITVTKILQALNI 240
QY 241 RIGRFSTGOKKARKKKTISLTQHEATDMSQSSGGRNVFVGVTSVSVIALLRAVKR 300
DB 241 RIGRFSTGOKKARKKKTISLTQHEATDMSQSSGGRNVFVGVTSVSVIALLRAVKR 300
QY 301 HRERERQKRVFMSLLIISTFLCWTPISVLNTTILCLGPSDLLVLCFLVMAYGTT 360
DB 301 HRERERQKRVFMSLLIISTFLCWTPISVLNTTILCLGPSDLLVLCFLVMAYGTT 360
QY 361 IFHPLLAFTRQKFKQKVLKSKMKRRVSVIYEADPLPNNAVIHSWIDPKRKKITTFEDSE 420
DB 361 IFHPLLAFTRQKFKQKVLKSKMKRRVSVIYEADPLPNNAVIHSWIDPKRKKITTFEDSE 420
QY 421 IREKRLVPQVWTD 433
DB 421 IREKRLVPQVWTD 433

RESULT 2
AAM50840
ID AAM50840 standard; Protein; 433 AA.
XX
AC AAM50840;
XX

DT 01-MAY-2002 (first entry)
XX
DE Cysteine protease-like protein.
XX
KW Cysteine protease; mouse; transgenic mouse; transgenic animal;
KW animal model; gene disruption; gene targeting; gene detection;
KW therapy; enzyme.
XX
OS Mus musculus.
XX
PN WO200206445-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US224202.
XX
PR 19-JUL-2000; 2000US-219168P.
PR 19-JUL-2000; 2000US-219171P.
PR 19-JUL-2000; 2000US-219175P.
PR 23-JUL-2000; 2000US-221455P.
PR 13-DEC-2000; 2000US-256212P.
XX
PA (DELT-) DELTAGEN INC.
XX
PI Brennan TJ, Allen KD;
XX
DR WPI; 2002-179787/23.
DR N-PSDB; ABA91703.
XX
PT Non-human transgenic animal useful as a disease model and for
PT identifying agents that modulate gene expression and function,
PT comprises a disruption in a targeted gene e.g. cysteine protease-like
PT gene -
XX
XX Example 1; Fig 1; 61pp; English.

CC The present sequence is that of the protein encoded by the cysteine
CC protease-like gene sequence given in ABA91703. The invention
CC provides non-human transgenic animals containing targeted gene
CC disruptions, including disruptions of the cysteine protease-like
CC gene. A claimed gene targeting construct comprises a first
CC polynucleotide sequence homologous to a target gene, a second
CC polynucleotide sequence homologous to the target gene, a selectable
CC marker and optionally a screening marker. A cell, especially a
CC murine embryonic stem cell, and a transgenic animal comprising a
CC disruption in a target gene are claimed. In the present case,
CC a cysteine protease-like-specific targeting construct having the
CC ability to disrupt or modify cysteine protease-like genes was
CC created using the targeting arms (homologous sequences) given
CC in ABA91704 and ABA91705. Transgenic mice were generated. The
CC cell- and animal-based systems are useful as models for disease
CC and for identifying agents that modulate gene expression and
CC function, and as potential treatments for various disease states
CC and disease conditions. Methods of treating diseases associated
CC with disrupted targeted gene expression or function comprise
CC detecting and replacing mutated target genes through gene therapy.

SQ Sequence 433 AA;

Query Match 100.0%; Score 2192; DB 23; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.1e-225;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCFSPILEINMQSESNTVRDDIDDINTNMYPQLSPYPLSFQVSLTGFLMLIEIVLGLGSLN 60
DB 1 MCFSPILEINMQSESNTVRDDIDDINTNMYPQLSPYPLSFQVSLTGFLMLIEIVLGLGSLN 60
QY 61 TVLVLYCMKSNLINSVNIITMNLHLVDVLCVGCIPLTIVILLLESNTALICCFHEA 120
DB 61 TVLVLYCMKSNLINSVNIITMNLHLVDVLCVGCIPLTIVILLLESNTALICCFHEA 120
QY 121 CVSPASVSTAINVFAITLDRYDISVKPANRLTMTGRAVLMISWIFSFLIPFTEVN 180
DB 121 CVSPASVSTAINVFAITLDRYDISVKPANRLTMTGRAVLMISWIFSFLIPFTEVN 180

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:56:53 ; Search time 40 Seconds
(without alignments)
1442.438 Million cell updates/sec

Title: US-09-845-721-2

Perfect score: 2192

Sequence: 1 MCFSP1E1NQSESNTVR.....ITFEDSEIRKRLVPQVTD 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 2192 | 100.0 | 433 | 21 AAY90622 | Human G protein-co |
| 2 | 2192 | 100.0 | 433 | 23 AAM50840 | Cysteine protease- |
| 3 | 2192 | 100.0 | 433 | 23 AAM47903 | Human GPR22. Homo |
| 4 | 2183 | 99.6 | 433 | 21 AAY90656 | Human mutant G pro |
| 5 | 2168 | 98.9 | 433 | 20 AAY03770 | Human G-protein co |
| 6 | 282 | 12.9 | 430 | 14 AAR40772 | Sequence encoded b |
| 7 | 282 | 12.9 | 430 | 22 AAB66625 | Guinea pig CCKA re |
| 8 | 282 | 12.9 | 450 | 22 AAB66626 | Guinea pig CCKA re |
| 9 | 266.5 | 12.2 | 444 | 14 AAR38890 | Sequence encoded b |
| 10 | 266.5 | 12.2 | 444 | 18 AAW21567 | LETO rat cholecyst |

| | | | | | |
|----|-------|------|-----|-------------|---------------------|
| 11 | 266.5 | 12.2 | 444 | 22 AAB66618 | Rat pancreatic CCK |
| 12 | 253 | 11.5 | 428 | 18 AAW29102 | Human peptide horm |
| 13 | 253 | 11.5 | 428 | 18 AAW29104 | Enhanced CCK-A/gas |
| 14 | 253 | 11.5 | 428 | 22 AAB66630 | Human CCK A recept |
| 15 | 239 | 10.9 | 371 | 17 AAB66630 | Mouse neuropeptide |
| 16 | 237.5 | 10.8 | 450 | 15 AAR53263 | M. matalensis CCK |
| 17 | 237.5 | 10.8 | 450 | 15 AAR53290 | Mastomys gastrin r |
| 18 | 229.5 | 10.5 | 584 | 22 AAU03215 | Fruit fly G protei |
| 19 | 226.5 | 10.3 | 451 | 14 AAR40771 | Sequence encoded b |
| 20 | 224 | 10.2 | 450 | 22 AAU03809 | G protein-coupled |
| 21 | 222.5 | 10.2 | 452 | 22 AAB66619 | Rat brain CCKB rec |
| 22 | 222 | 10.1 | 422 | 22 AAB70249 | HTR1A protein. Ho |
| 23 | 219.5 | 10.0 | 411 | 17 AAW03012 | Neuropeptide Y rec |
| 24 | 219 | 10.0 | 447 | 14 AAR40774 | Sequence encoded b |
| 25 | 219 | 10.0 | 447 | 22 AAB66629 | Human CCK B recept |
| 26 | 218.5 | 10.0 | 370 | 15 AAR57066 | Murine serotonin |
| 27 | 218 | 9.9 | 447 | 15 AAR53264 | Human CCK B recept |
| 28 | 218 | 9.9 | 447 | 15 AAR57738 | Human cholecystoki |
| 29 | 218 | 9.9 | 447 | 18 AAW29101 | Human peptide horm |
| 30 | 217 | 9.9 | 447 | 21 AAB02846 | Human G protein co |
| 31 | 215 | 9.8 | 447 | 21 AAB02852 | Human G protein co |
| 32 | 215 | 9.8 | 447 | 21 AAB02855 | Human G protein co |
| 33 | 214 | 9.8 | 453 | 14 AAR40773 | Sequence encoded b |
| 34 | 214 | 9.8 | 453 | 22 AAB66627 | Guinea pig CCKB re |
| 35 | 213.5 | 9.7 | 357 | 15 AAR58685 | Rat RECL17 serotoni |
| 36 | 213.5 | 9.7 | 407 | 16 AAR80749 | Canine cholecystok |
| 37 | 213.5 | 9.7 | 407 | 17 AAR92291 | Canine cholecystok |
| 38 | 213.5 | 9.7 | 407 | 17 AAR88460 | Canine cholecystok |
| 39 | 213.5 | 9.7 | 421 | 22 AAB56316 | Non-endogenous hum |
| 40 | 213 | 9.7 | 453 | 22 AAB66628 | Canine gastrin rec |
| 41 | 212.5 | 9.7 | 460 | 22 AAB61970 | Rat HCTR2 polypep |
| 42 | 212 | 9.7 | 447 | 14 AAR41678 | Human cholecystoki |
| 43 | 211.5 | 9.6 | 357 | 22 ABB56327 | Non-endogenous hum |
| 44 | 211.5 | 9.6 | 478 | 21 AAB36476 | Human SNORF36a rec |
| 45 | 211.5 | 9.6 | 515 | 19 AAW77106 | Rat alpha18-adrene |

ALIGNMENTS

RESULT 1

| | |
|---------|--|
| AA90622 | AA90622 standard; Protein; 433 AA. |
| ID | AA90622 standard; Protein; 433 AA. |
| AC | AA90622; |
| XX | |
| DT | 21-AUG-2000 (first entry) |
| XX | |
| DE | Human G protein-coupled receptor GPR22. |
| XX | |
| KW | G protein-coupled receptor; GPCR; constitutively active; |
| KW | intracellular loop 3; transmembrane domain 6; drug screening; |
| KW | agonist; antagonist. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200022129-A1. |
| XX | |
| PD | 20-APR-2000. |
| XX | |
| PF | 12-OCT-1999; 99WO-US23938. |
| XX | |
| PR | 13-OCT-1998; 98US-0170496. |
| XX | |
| PA | (AREN-) ARENA PHARM INC. |
| XX | |
| PI | Behan DP, Chalmers DT, Liaw CW; |
| XX | |
| DR | WPI; 2000-329165/28. |
| XX | |
| PT | N-PSDB; AAA30613. |
| XX | |
| PT | Non-endogenous constitutively activated human G protein-coupled |
| PT | receptors, useful for identifying agonists for use as pharmaceutical |